MANAGING UNCERTAINTY IN ENGINEERING DESIGN USING IMPRECISE PROBABILITIES AND PRINCIPLES OF INFORMATION ECONOMICS

A Dissertation
Presented to
The Academic Faculty

by

Jason Matthew Aughenbaugh

In Partial Fulfillment
Of the requirements for the Degree
Doctor of Philosophy in the George W. Woodruff School of Mechanical Engineering

Georgia Institute of Technology

August, 2006

MANAGING UNCERTAINTY IN ENGINEERING DESIGN USING IMPRECISE PROBABILITIES AND PRINCIPLES OF INFORMATION ECONOMICS

Approved by:

Dr. Christiaan J. J. Paredis, Chair
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Leon McGinnis
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Bert Bras
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Mark Ferguson
College of Management
Georgia Institute of Technology

Dr. Ye-Hwa Chen
School of Mechanical Engineering
Georgia Institute of Technology

Date Approved: June 8, 2006

ACKNOWLEDGEMENTS

Above all others, I would like to thank my family for their never-ending support and dedication. My mother and father have always put me first in their lives, and I would not be the person that I am today without their nurturing, guidance, discipline, and encouragement. I would also like to thank my grandparents, who throughout their lives worked to give my parents great opportunities and also gave back to society through teaching, engineering, and community service. They have served as role models for hard work and dedication.

Academically, I am very grateful to my advisor, Dr. Chris Paredis. He simultaneously gave me the academic freedom to pursue a topic of interest to me and the necessary guidance to realize my goals. I would particularly like to acknowledge his willingness to debate issues and to engage in productive dialog. This intellectual exchange contributed greatly to my research contributions and my enjoyment of my graduate studies.

I am also very grateful to the National Science Foundation for my Graduate Research Fellowship. This financial support enabled me to pursue a novel research topic and to push myself to make a significant contribution to the engineering design community. I would also like to thank the Georgia Institute of Technology and Tau Beta Pi for additional fellowship support during my graduate studies.

I would also like to thank the members of the Systems Realization Laboratory at Georgia Tech. This group provides a unique environment of support, both socially and academically. I would like to specifically thank Rich Malak, Jay Ling, Morgan Bruns, Scott Duncan, Steve Rekuc, and Tarun Rathnam for their valuable feedback and collaboration on my work. They have also been great friends and added enjoyment to my life outside of the office. I would also like to note Chris Williams and Matt Chamberlain, who always made tailgating more fun and sports more competitive. I thank Dr. Farrokh Mistree for offering advice, providing inspiration, and challenging me to define and achieve my goals.

I would like to thank my committee—Dr. Bras, Dr. Chen, Dr. McGinnis, and Dr. Ferguson—for their time and feedback.

I also owe many thanks to various professors, teachers, friends throughout the years. There is more to life than research, and without my friends, my life would be unbalanced. I would like to thank my great friends and former roommates Eric Anderson—who was always up for something fun and for academic cooperation—and Deepak Hedge—who kept me thinking about stimulating topics that do not involve engineering design. These two greatly enriched my Georgia Tech experience. Finally, I would like to thank the Atlanta Flying Disc Club and its members, who have kept me exercising, enjoying Atlanta's weather, and having fun times during the last two and a half years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY

aleatory uncertainty Uncertainty that arises from a random process.

certainty The condition of knowing everything necessary to choose the course of action whose outcome is most preferred.

consistent region The region in the hyperspace of imprecise quantities that is consistent with the available information.

decision alternative A specific option for a specific decision.

design alternative One of the possible complete product design specifications.

decision robustness, sensitivity analysis for The goal in sensitivity analysis of determining whether the current decision is robust, given the DM's state of incomplete information.

ecosphere In a life cycle analysis, modeling of changes to the environment.

epistemic uncertainty Uncertainty resulting primarily from a lack of knowledge.

frequentist interpretation An interpretation of probability theory in which a probability represents the limit of the ratio of times that one outcome occurs compared to the total number of outcomes in an endless series of identical trials

general p-box A p-box that contains all non-decreasing distributions that fall entirely between the bounding distributions.

imprecision The gap (if any) between the present state of information and a state of precise information. More generically, the quality or state not being exactly or sharply defined or stated.

information economics A set of principles derived from standard economic principles and applied to the collection and modeling of information in engineering design.

information prioritization, sensitivity analysis for A goal of sensitivity analysis in which a DM seeks to understand what the most useful information to collect is, given the existing lack of robustness in the decision

irreducible uncertainty Uncertainty associated with things that are inherently unknowable.

operational definition The operational definition of a quantity is a set of operations, or a procedure, by which that quantity can be measured

parameterized p-box A p-box that contains only distributions specified by a distribution type and intervals of the parameters of the distribution.

reducible uncertainty Uncertainty associated with things that are knowable but are currently unknown.

state of precise information The state of having acquired all information about a particular model of irreducible uncertainty available at any price.

subjective interpretation An interpretation of probability theory in which a probability represents an individuals beliefs about an uncertain quantity in terms of the individual's willingness to enter into gambles at different prices

technosphere In a life cycle analysis, description of the product and its life cycle and an inventory of loads (e.g. emissions)

uncertain parameter A parameter, or quantity, about which a decision maker has uncertainty. It is more general than a random variable because the uncertainty need not be representing using probability theory.

uncertainty The gap between was is currently known and certainty

valuesphere In a life cycle analysis, modeling of the perceived seriousness or importance of changes to the environment

LIST OF SYMBOLS

μ	mean of an uncertain quantity
$\hat{\mu}$	DM's best estimate of the mean of an uncertain quantity
Σ	a set of statistical data samples
σ^2	variance of an uncertain quantity
$\hat{\sigma}^2$	DM's best estimate of the variance of an uncertain quantity
$A = \{a_1, ..., a_m\}$	a set of alternative actions for a DM
a	one of the possible actions or alternatives available to a DM
$\underline{E}[X]$	lower bound on the expected value of an uncertain quantity X
$\overline{E}[X]$	upper bound on the expected value of an uncertain quantity X
$E_F[X]$	the mathematical expectation of uncertain quantity X taken over the probability distribution with cumulative probability distribution function $F_X(x)$
$E_f[X]$	the mathematical expectation of uncertain quantity X taken over the probability distribution with probability mass function $f_X(x)$
$F_X(x)$	the probability that the value of the uncertain quantity X is less than some set value x. In general, $F_X(x) = P(X \leq x)$. For a continuous uncertain quantity with probability mass function $p(x)$
$\overline{F}_X(x)$	upper bound on the CDF for an uncertain quantity X
$\underline{F}_X(x)$	lower bound on the CDF for an uncertain quantity X

$f_X(x)$	probability mass function for an uncertain quantity X
$G(I)$	the set of all consequences that are consistent with the available set of information I
$g(a_i, s_j)$	the consequence of taking action a_i given that state s_j of the world occurs
I	the set of information available to a DM, which includes information about possible actions, the states of the world, and the DM's own beliefs and preferences
$P(I)$	the set of probabilities that are consistent with the available set of information I
$\underline{P}(X)$	lower probability or prevision for gamble X
$\overline{P}(X)$	upper probability of prevision for gamble X
$p_j({}_i x)$	a particular probability density or probability mass function or distribution for uncertain quantity ${}_i x$
$S = \{s_1, \ldots, s_r\}$	a set of alternative states of the world that can be realized
s	(1) a state of the world
	(2) the sample standard deviation
$U(I)$	the set of all utility functions consistent with the available information I
$\tilde{V}(n+1)$	Estimated value of the $(n+1)^{st}$ data sample in an information collection problem
${}_j X$ or just X	a uncertain quantity or a particular gamble
${}_j X_i$	a particular realization i of uncertain quantity ${}_j X$
$\underline{x}$	Lower bound on an uncertain quantity

$\overline{x}$ Upper bound on an uncertain quantity

$\boxed{X}$ A general p-box for uncertain quantity X.

$$\boxed{X} = \left\{ F_X(x) : \forall x \in \mathbb{R},\ \underline{F}_X(x) \le F_X(x) \le \overline{F}_X(x) \right\}$$

$\boxed{X}^P$ a parameterized p-box for uncertain quantity X. for example, if the set of distributions is limited to normal distribution:

$$\boxed{X}^P = \left\{ F_X(x;\mu,\sigma) = \Phi_{\mu,\sigma}(x) : \mu \in \left[\underline{\mu},\overline{\mu}\right], \sigma \in \left[\underline{\sigma},\overline{\sigma}\right] \right\}$$

LIST OF ABBREVIATIONS

CDF	cumulative distribution function
DASA	Decision analysis with sensitivity analysis
DBC	dependency bounds convolution
DLS	double loop sampling
DM	decision maker
EBDM	environmentally benign design and manufacture
OPS	optimized parameter sampling
PBA	probability bounds analysis
PCS	p-box convolution sampling
PDF	probability density function
PEC	plastic easy-change, a type of oil filter
SBCE	Set-Based Concurrent Engineering
SEC	steel easy-change, a type of oil filter
SEU	subjective expected utility
TASO	take-apart spin-on, a type of oil filter

SUMMARY

The engineering design community recognizes that an essential part of the design process is decision making. Each decision consists of two main phases—problem formulation and problem solution. Because decisions generally are made under uncertainty, engineers need appropriate methods for modeling and managing uncertainty. Existing literature focuses on modeling uncertainty using precisely known probabilities. The objective of this thesis is to investigate and develop alternative methods for managing uncertainty during the formulation phase of engineering design decisions, focusing on situations in which probabilities are not known precisely.

Two important characteristics of uncertainty in the context of engineering design are imprecision and irreducible uncertainty. In order to model both of these characteristics, it is valuable to use probabilities that are most generally imprecise and subjective. These imprecise probabilities generalize traditional, precise probabilities; when the available information is extensive, imprecise probabilities reduce to precise probabilities. However, when information is scarce, they more accurately represent a decision-maker's uncertainty.

An approach for comparing the practical value of different uncertainty models is developed. The approach examines the value of a model using the principles of information economics: value equals benefits minus costs. The benefits of a model are measured in terms of the quality of the product that results from the design process. Costs are measured not only in terms of direct design costs, but also the costs of creating and using the model.

Using this approach, the practical value of using an uncertainty model that explicitly recognizes both imprecision and irreducible uncertainty is demonstrated in the context of a high-risk engineering design example in which the decision-maker has few statistical samples to support the decision. It is also shown that a particular imprecise probability model called probability bounds analysis generalizes sensitivity analysis, a process of identifying whether a particular decision is robust given the decision-maker's lack of information. An approach for bounding the value of future statistical data samples while collecting information to support design decisions is developed, and specific policies for making decisions in the presence of imprecise information are examined in the context of engineering.

CHAPTER 1:

INTRODUCTION

During the engineering design process, a designer's actions are constrained by limited resources. Consequently, the information available for guiding analysis and decision-making is generally incomplete, and the designer must make design decisions under a state of uncertainty. The majority of existing design research has focused on solving design problems—that is, choosing a specific design. Much less attention has been given to the process of formulating design problems, which includes collecting relevant information and transforming this information into a form that can be used to support decision making.

Some research advancements, such as improved computer modeling and increased computing power, are complementary to both the formulation and solution phases of design problems. For example, increases in computing power enable engineers to run analysis models that were previously infeasible. Other research has focused on computer algorithms, such as approaches to optimization. Improved analysis and solution methods help designers solve the formulated problem more accurately than they could with previous methods, but the solution to the formulated problem will only match the true solution if the problem is formulated appropriately.

Despite the significant advancements in engineering research with respect to solving design problems, there has been much less attention paid to the fundamental problem of formulating the design problem, including the crucial recognition of what is known (i.e. information) and what is not known (i.e. uncertainty). Specifically, the underlying models of uncertainty used by practicing engineers and researchers have remained mostly the same for decades; traditional practice uses perfectly known, precise probability distributions.

In this dissertation, it is argued that there is more to uncertainty than just perfectly known probabilities. For example, often one goal of information collection is to learn about the probabilities of events and the dependencies between events, such as the experiment of drawing samples from a random process to characterize its mean and variance. The goal of this experiment is to increase knowledge about the probabilities, which implies that there can be lack of knowledge, or uncertainty, about probabilities. If there is uncertainty about probabilities, can probabilities still be the most general representation of uncertainty?

The goal of this thesis is to investigate methods for managing uncertainty and information during the formulation phase of engineering design decisions, focusing on situations in which engineers do not have complete information. The main motivating question of the dissertation is:

How should engineering designers manage information to support decision making under uncertainty?

This question will be answered from two perspectives:

- A *theoretical* perspective—identifying the internal consistency and applicability of representations to design problems
- A *practical* perspective—given the benefits and costs associated with different methods, identifying the method that yields the highest overall economic value to the design process.

The proposed answer to this question is that engineers need to represent uncertainty using methods that go beyond precise probabilities. These methods should generalize from probability, meaning that they should include probability as a special case. At the same time, these methods should recognize that often engineers lack perfect knowledge about the problem. It is also argued that uncertainty due to a lack of knowledge should be represented distinctly from inherent randomness, because otherwise lack of knowledge and randomness become intertwined and indistinguishable. This confounding of two different things makes analysis difficult, including complicating the management of information collection that seeks to reduce uncertainty.

The remainder of this chapter establishes the context of the research (Sections 1.1-1.4), identifies the secondary motivating questions addressed in the dissertation (Section 1.5), and provides an overview of the organization of the remaining chapters (Section 1.6).

1.1 The engineering design context

Design is a process of converting information about customer interests and requirements into a specification of a product. This process involves searching through a very large, unstructured space of solutions (Tong and Sriram 1992) based on vague and uncertain knowledge about possible solution alternatives (Gupta and Xu 2002), their physical behavior (Aughenbaugh and Paredis 2004), their cost (Garvey 1999), and the decision maker's (DM's) preferences (Kirkwood and Sarin 1985, Otto and Antonsson 1992, Carnahan, et al. 1994, Seidenfeld, et al. 1995). In order to guide engineers through this process, several approaches have been developed. One approach to this process is the systematic design method described by Pahl and Beitz (1996).

1.1.1 Systematic design

In systematic design (Pahl and Beitz 1996), the design process is broken into four main phases, as summarized in Table 1.1. In the product planning and clarification of task phase, a need for a product is determined and described. Product planning is mostly in the domain of corporate strategy and marketing; a company's situation and market condition are analyzed, profitable product ideas sought, and a product proposal made. The next step is to clarify the task by refining the product proposal and creating a detailed requirements list for the product. These requirements tell engineers what a product should be, should not be, and what it must be (at a minimum) in order to be successful. Once a list of requirements and objectives is created, conceptual design can begin.

The conceptual design phase takes the list of requirements and objectives and determines the principle solution structures to be pursued in embodiment design. To some, this is where traditional engineering begins. First, designers distill the problem down to its core, asking *what are we really trying to build*. Then they identify what functions (e.g. in a car design, functions such as move person, protect person, monitor performance) the design must perform and how these functions interact at a high level, such as transfers of energy, mass, and signals. All of this information is combined into a function structure.

Next, designers seek to enumerate possible physical implementations, or working principles, for each function. For example, three working principles for the function *mark a piece of paper* could be *deposit material by friction* (e.g. a pencil), *melt material onto paper* (e.g. laser jet printing), or *burn away material* (e.g. scorching the paper with a laser). Since in general there are multiple functions, each with multiple working principles, they can be combined into an overall product in many different ways, or solution variants. Finally, these solution variants must be evaluated and a principal solution concept chosen. This concept forms the foundation for embodiment design.

Table 1.1. Systematic design phases

Phase	Main tasks
1. Planning and clarifying the task	Investigation into the economic and technical viability of creating a given product, and the definition of the exact requirements of a system and the criteria surrounding its functioning.
2. Conceptual design	Development of function structure and the evaluation of different solution variants to this problem.
3. Embodiment design	Conversion of a conceptual working structure to a specification of layout.
4. Detail design	Finalization of the design and production details.

In embodiment design, designers develop the design concept in more detail by considering additional technical and economic criteria. Essentially, embodiment design takes the working principles and concepts developed in conceptual design and develops an actual design specification, at which point detail design can lead directly into production. During detail design, the arrangement, dimensions, materials, and production methods of all parts of the product are finalized and documented.

1.1.2 Partitioning the design problem

Complex problems can rarely, if ever, be solved globally in one step. Most products have reached a level of complexity at which it is infeasible for one engineer, or even engineers from a single discipline, to design them completely. Instead, the design problem must be broken down into smaller chunks that are designed by separate design teams. The solutions to these sub-problems are then synthesized and integrated into a complete design for the overall system. Systematic design is an appropriate approach for designing a product at one level of detail, but it does not address this higher-level process of decomposing a system into subsystems, concurrently designing subsystems, and subsequently integrating subsystem designs into the overall system. A holistic, hierarchical decomposition approach to the design process that addresses these problems is provided by systems engineering (Forsberg and Mooz 1992, Buede 2000, Forsberg, et al. 2000, Blanchard 2004). Although this dissertation will not address systems engineering formally (see Aughenbaugh and Paredis (2004) for a discussion of modeling, simulation, and uncertainty in systems engineering), it is useful to consider the consequences of decomposing the design process.

When the design process is decomposed, it becomes recursive—the overall design process is a sequence of design sub-problems. For example, consider the design of a car. A car can be broken down into many subsystems (such as engine, drivetrain, wheels, chassis, and so on), and each of these subsystems can be broken down into smaller subsystems, as cartooned in Figure 1.1. In many cases, a different team of engineers will perform the embodiment of each subsystem. Teams may also work on sub-problems concurrently, rather than sequentially. For example, one team may be designing the drivetrain while another team is designing the engine.

When a team is formed to design the engine, its members first must clarify their task by using their technical expertise to elaborate on the requirements. For example, a particular engine concept is one of the working solutions from the conceptual design (phase 2 of systematic design) of the car, as shown in Figure 1.2. Part of this engine design process is subdividing the engine into its subsystems, such as the fuel intake, and so on down to the smallest component of the system. This design process is challenging because the performance of the overall system may be a function of the interactions between sub-systems. Thus, the decisions of one team depend on future decisions and on decisions made concurrently by other design teams. The decisions outside of the control of one group are sources of uncertainty to that group's decision. The importance of decisions in the design process, including decisions under uncertainty, has lead many researches to adopt a perspective of decision-based design.

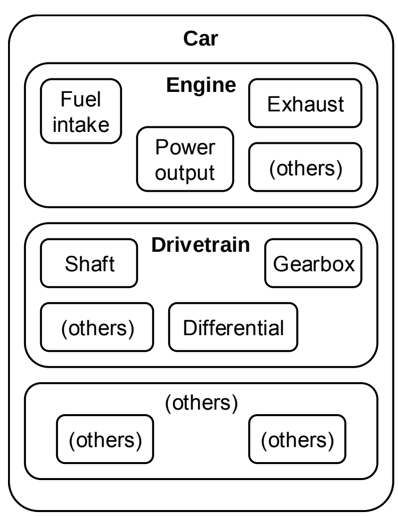


Figure 1.1. Subsystems of a car

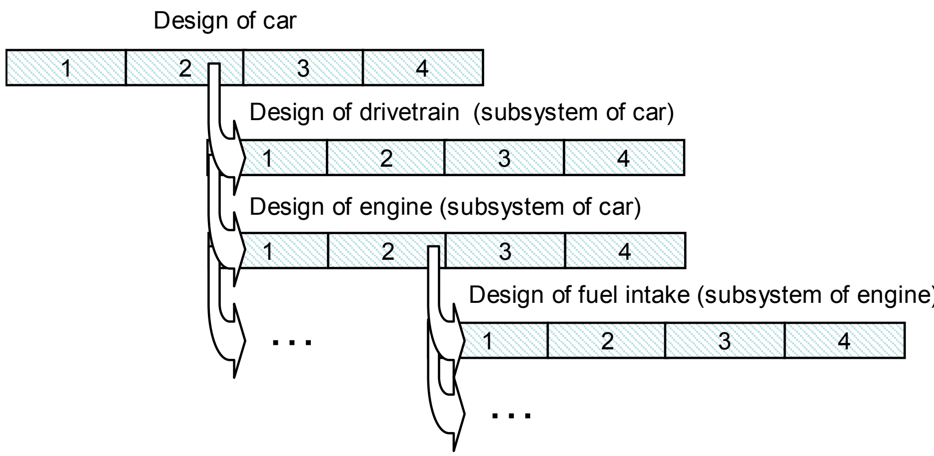


Figure 1.2. Recursive systematic design process, design phases numbered 1-4

1.1.3 Decision-based design

Independent of the design method that an organization adopts, designers repetitively must identify problems, search for solutions, evaluate solutions, and choose an action or design. Inspired by this process, decision-based design recognizes that the principal role of an engineer in the design process is to make decisions (Mistree, et al. 1990, Hazelrigg 1998, Marston, et al. 2000). This paradigm shifts the emphasis of design research to decision making; the motivation is that one way to improve the design process is to enable engineers to make better decisions.

Engineers have finite resources with which to conduct the design process. Consequently, they cannot study every detail of every subsystem extensively. Decisions often are guided with approximate models, expert opinion, rules of thumb, and even pure intuition. In general, it is very difficult for the right person to have the right information available at the right time in a format that he or she can comprehend (Cooper 2003). One goal of decision-based design is to support decisions with formal methods that make the most out of the available information and resources. The idea is that a final design can only be as good as the decisions that led to it, so the decisions need to be as good as possible. The open question is what formal methods are most appropriate for engineering design. This question is next viewed in the more specific context of simulation-based design.

1.1.4 Simulation-based design

As already noted, the partitioning of the design problem into sub-problems results in a sequence of decisions (for simplicity, concurrent decisions by multiple decision makers are ignored), of which one is illustrated in Figure 1.3. In this grossly abstracted model of the design process, a designer or decision maker (hereafter abbreviated as DM), has identified two decision alternatives. The DM performs multiple simulations (S_i) or other analyses (A_i), including eliciting expert opinion, to study the performance of the alternatives in various environmental factors. Performance attributes are then combined or weighted according to the DM's preferences, perhaps according to utility theory(von Neumann and Morgenstern 1944, von Neumann and Morgenstern 1980, Keeney and Raiffa 1993). Finally, the most preferred alternative is selected, or alternatively, when there are more than two decisions alternatives, the DM can proceed by selecting a set of alternatives and subsequently eliminating the inferior solutions (Rekuc, et al. 2006), a topic considered in detail in Chapter 8.

The abstract model in Figure 1.3 is useful because even in its simple form, it portrays the complexity of the design process. Every component in the figure—

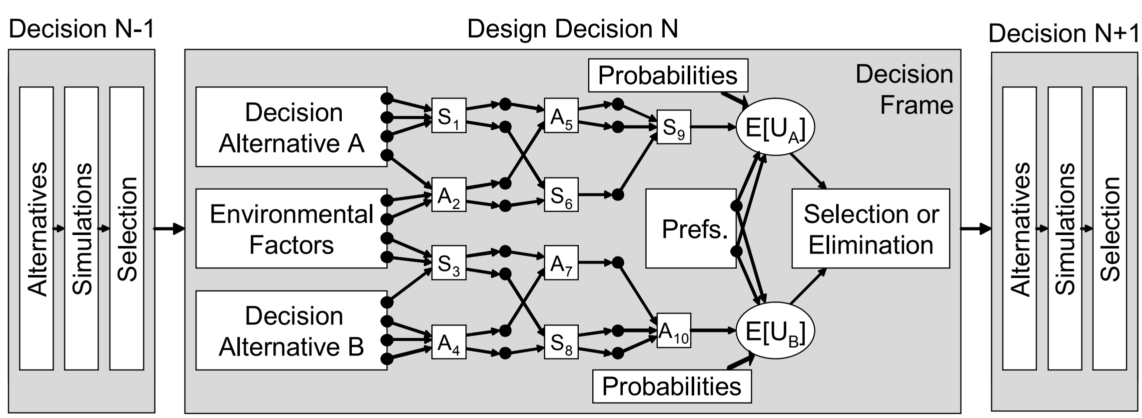


Figure 1.3. Abstraction of a sequential decision process in simulation-based design.

alternatives, environmental factors, simulations, analysis, preferences, and probabilities—introduce uncertainty into the decision. In order to make a good decision, a DM should recognize and account for these uncertainties when making the decision. The motivating research question in this dissertation asks how to best manage information in this process. Because information and uncertainty are duals (see Section 1.3), it also asks the question of how these uncertainties can best be propagated through the analysis process and then incorporated into the decisions, a question that involves both the formulation and solution of the decision problem.

1.1.5 Decision problem formulation and solution

Decisions have two phases—problem formulation and problem solution. The decision formulation phase involves an important sub-decision problem, namely, how much information to collect in support of decision making, that is, the solution of the decision problem.

The basic elements of formulating a decision problem and collecting information can be summarized in five steps, adapted from Kmietowicz and Pearman (1981):

1. Identify an exhaustive set of mutually exclusive decision alternatives, which could include ranges of continuous parameters
2. Identify an exhaustive set of mutually exclusive alternative states of the world, or alternatively define the appropriate continuous variables that capture the state
3. Predict the payoff of every decision alternative in every state of the world
4. Assign probabilities to each state of the world (including assigning probability mass functions to any continuous states), or acknowledge that such information is not available
5. Select the criteria for evaluating alternatives

Only after formulating the problem can a DM attempt to solve the problem by identifying the decision alternative that is most preferred. In general, engineers must complete the formulation process within significant constraints, such as deadlines, monetary budgets, and bounded rationality—a phrase coined by Herbert Simon (1947) that refers to the inherent bounds on human thinking. It can thus be difficult or impossible to carry out all five steps of the formulation process in their entirety (Murphy, et al. 2005) [see Section 2.4 for more on particular causes of a DM's lack of information in design]. Without such constraints, it would be possible to make use of a single objective decision making, traditional probability theory, and von Neumann-Morgenstern utility theory (1944), as recommended by Hazelrigg (1996) and others. But what should be done when the constraints do exist?

Despite the existence of such constraints in practice, the problem formulation phase of the decision problem has received much less attention than the problem solution phase. Recent literature (Bradley and Agogino 1994, Gupta and Xu 2002, Radhakrishnan and McAdams 2005) acknowledges the reality of resource constraints and the impracticality of exhaustive analysis, but it presents few alternatives for problem formulation. Gupta and Xu (2002) note the impossibility, given time and budgetary constraints, of exploring all possible design alternative payoffs (steps 1-3), and they identify significant tradeoffs in the number of alternatives considered, but they present no guidance for actually managing the design process in this aspect. Radhakrishnan and McAdams (2005) analyze the cost-benefit trade-offs in selecting models of various levels of abstraction in engineering design and present a framework in which a designer can reason about model uncertainty, but the designer is left with little guidance in estimating actual value of information from different models. Bradley and Agogino (1994) develop a decision-analytic approach to assist designers in cost-benefit analysis of resource expenditures using precisely characterized probability distributions to guide and prioritize information collection, but they do not explain how to estimate these distributions.

Other work has focused on evaluation, Step 5. Engineers have developed or adopted various methods to support design decisions under uncertainty, such as statistical decision theory (Pratt, et al. 1995) and utility theory (von Neumann and Morgenstern 1944), decision analysis (Howard 1968, 1988a, 1988b), safety factors (Elishakoff 2004), probabilistic risk assessments (Bedford and Cooke 2001), reliability based design optimization (Mourelatos and Liang 2004), and robust design (Byrne and Taguchi 1987, Taguchi 1987, Allen, et al. 2006). Each of these methods requires the designer to formalize preferences in some way. The process of eliciting and formalizing preferences is not necessarily trivial, and there can be significant imprecision in utility functions (Smith 1961, Aumann 1962, 1964, Kirkwood and Sarin 1985, Weber 1987, Thurston 1990a, 1990b, Otto and Antonsson 1991, Thurston, et al. 1991, Otto and Antonsson 1992, Carnahan, et al. 1994, Antonsson and Otto 1995, Seidenfeld, et al. 1995), but it is not the focus of this research. In this dissertation, it is assumed that designers can capture their preferences accurately and completely using a single utility function. The combined problem of incompletely known preferences and incompletely known future states of the world (e.g., incompletely known probabilities), is left for future work.

In addition to the above limitations, most methods also suffer from the lingering assumption of the existence of clearly defined probabilities. That is, in most engineering design methods, it is assumed that Step 4 can be completed with little difficulty; the possibility that the probabilities are unknown is often ignored, even though decision theory has long acknowledged that they are not always known (Knight 1921). Knight specifically distinguishing decisions in which probabilities are known—decisions under risk—and cases in which probabilities are not known—decisions under uncertainty. This terminology is not used in this dissertation, but the concepts are closely related to the distinctions made about characteristics of uncertainty in Chapter 2.

From where do designers get probabilities in Step 4? How should these probabilities be represented? Are probabilities even appropriate for representing uncertainty in engineering design? These questions are exciting from both a theoretical/philosophical perspective—what is uncertainty?—and from a practical perspective—what method is most valuable to the designer?

1.2 Information economics

While philosophical questions are interesting to ponder, this dissertation goes beyond them by also tackling practical issues—in practice, how should engineers overcome resource constraints and uncertainty in the design process? This question is answered not only from a mathematical perspective, but also from a practical and economic perspective. Different formalisms for uncertainty have different assumptions, different methods, and different costs associated with their use. Part of managing uncertainty in engineering design is choosing the "right" representation. Another part is collecting the "right" amount of information to make a decision. So how can "right" be defined? Extending the principles of information economics (Marschak 1974), an engineer should choose the methods, models, and information that accord the greatest net value—benefit minus cost—to the entire design process and product lifecycle. By focusing on the practical value, the usefulness of the methods can be evaluated directly.

In most of this dissertation, information economics is used as a motivating and guiding principle rather than a specific method of analysis. Restated from above, this basic principle is that an engineer should only take a course of action—whether purchasing information, performing experiments, or using a particular model—if the benefits of that action outweigh its costs. The prediction of the benefits of a model is a difficult problem because the results and impact of a particular model are uncertain. A more detailed discussion of the application of information economics to information collection in engineering design, including a proposed method for estimating the value of additional information collection, is presented in Chapter 9. However, the practical, economic principle of cost-benefit analysis is used throughout, especially in Chapter 5 where the practical values of two models of uncertainty are compared.

1.3 Information and uncertainty modeling

There is a definite relationship between information and uncertainty; the more information one has, the less uncertainty exists. As such, information management and uncertainty management are intimately linked. Most of this dissertation approaches the problem from the perspective of uncertainty. The motivation is that in order to make good design decisions, a DM must recognize and account for the uncertainty that actually exists because overlooking this uncertainty can result in under-designed systems and catastrophic failure, or over-designed systems that require more resources than necessary.

An engineer recognizes uncertainty by *modeling* uncertainty. Like most aspects of the real world, the true state of information and uncertainty is often too complex to deal with exactly, so a DM instead builds a manageable model. To be useful, this model should contain the most important aspects of the state of the world for the given problem. To be efficient, the model should not contain much unnecessary information, as this increases the cost of building and using the model but provides no benefit. In this context, the subject of this dissertation is the process of building a theoretically justifiable and practically useful model of uncertainty in engineering design.

It is useful throughout this dissertation, especially when studying the characteristics of uncertainty (Chapter 2) and then exploring possible models of uncertainty (Chapter 3), to keep in mind a quote from Edward de Bono :

> The purpose of science is not to analyse or describe but to make useful models of the world. A model is useful if it allows us to get use out of it.

Applied to the topic of uncertainty, this quote can be taken to mean that the goal of the scientific study of uncertainty does not need to be absolute truths, but only models that help designers make better designs. This emphasizes the practical perspective of information and uncertainty modeling. However, the theoretical part of the problem, which for uncertainty modeling is somewhat philosophical, is not altogether irrelevant; so it is addressed in Chapter 2. The remainder of this section discusses practical aspects of the problem.

As mentioned in Section 1.1.5, engineers currently recognize uncertainty in a variety of ways. The most basic way is through the use of safety factors (Elishakoff 2004). This simple approach presents a useful context in which to discuss uncertainty modeling. When engineers apply safety factors in design, they are stating that they know that there is uncertainty in the analysis models. Because of this uncertainty and the assumptions that they adopt during analysis, the results of their analyses differ from the true state of the world.

For example, if an engineer is building a pressure vessel with material yield strength σ_y, the requirement to avoid failure is that σ_y exceeds the maximum stress in the pressure vessel, $\sigma_{\max}$. That is: $\sigma_y > \sigma_{\max}$. In a safety factor approach, the engineer employs a safety factor $SF > 1$ with his or her best point estimates $\tilde{\sigma}_y$ and $\tilde{\sigma}_{\max}$ of the true σ_y and $\sigma_{\max}$, respectively. The engineer then designs the pressure vessel such that $\tilde{\sigma}_y > SF \times \tilde{\sigma}_{\max}$. The engineer hopes that by designing the pressure vessel with a safety margin around the estimates, the true yield strength σ_y (which may be much less than $\tilde{\sigma}_y$) will exceed the true $\sigma_{\max}$ (which may be much larger than $\tilde{\sigma}_{\max}$).

What does the use of a safety factor say about the engineer's model of uncertainty? Assuming that the engineer is striving for 100% reliability (an uncommon goal, but one with clear meaning), then it essentially says that the uncertainty in the estimates $\tilde{\sigma}_y$ and $\tilde{\sigma}_{\max}$ is small enough that the SF fully compensates for it. Mathematically, this is equivalent to Equation (1.1), where $(\tilde{\sigma}_{\max} + \varepsilon_1)$ represents the absolutely largest stress that could occur in the pressure vessel walls and where $(\tilde{\sigma}_y - \varepsilon_2)$ represents the absolutely smallest yield strength of the pressure vessel that the engineer believes is possible.

$$(\tilde{\sigma}_y - \varepsilon_2) > SF \cdot (\tilde{\sigma}_{\max} + \varepsilon_1) \tag{1.1}$$

This notation emphasizes that the DM is making two estimates—an estimate of the largest maximum stress in the walls of the pressure vessel and an estimate of the smallest yield strength of the material. In order for the design to be reliable, a safety factor is chosen such that Equation (1.1) holds. However, since ε_1 and ε_2 are unknown, the choice of an appropriate safety factor is a challenge.

Is Equation (1.1) a sufficient model of the uncertainty? The answer depends on an information economic analysis. In order to be sure that the inequality in Equation (1.1) holds, an engineer may have to choose a very large safety factor because ε_1 and ε_2 are uncertain. A larger safety factor in general corresponds to a more costly design, such as a pressure vessel with thicker walls. If this cost is small compared to the benefits of guaranteeing reliability, it may be a valuable tradeoff. For example, consider moving to a probabilistic model of uncertainty. The construction of a probabilistic model requires significantly more data than a basic safety factor approach (Elishakoff 2004). The acquisition of this data may take considerable time and money, in which case the costs of collecting the information that the probabilistic model requires may exceed the costs of over-design from the safety factor approach.

The movement from a safety factor model to a probabilistic model of uncertainty represents the general situation of comparing models. Different models generally are based on different assumptions and require different information. At one level, it makes sense to use a model that captures exactly the available information, meaning a model that fully represents what is known without requiring additional assumptions. This model would be as close to reality as possible. However, this model could be very complex and

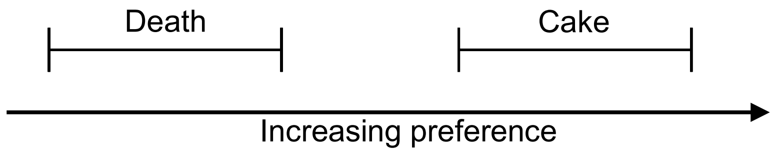


Figure 1.4. Imprecise model example: cake or death.

difficult to use to make calculations and decisions. Therefore, at another level it may not be economically valuable to use a complicated model.

Hazelrigg (2003) has used a skit by Eddie Izzard to illustrate this point. In an extension of this example, a person is presented with a choice between cake or death (i.e., eating a piece of cake or dying). For most people, even the very abstract model of "cake"—the person is not told the flavor, freshness, size, or any other detail of the cake—is sufficient for them to choose cake over death, even though the model of death—the person is not told how he or she will die—is also very abstract.

In Figure 1.4, a DM's preferences for cake and death are shown as intervals. Intervals are used because no single value of preference can be associated with the rough models of cake and death. For example, a person may really like chocolate cake with vanilla icing but hate white cake with peanut butter icing. Each type of cake has a certain preference associated with it, and together the class of all "cakes" forms an interval of preferences.

Even without specific information, the person can conclude that cake is preferred to death, since the entire interval for cake is at a higher level of preference than the entire interval for death (i.e., the interval for cake is entirely to the right of the interval for death in the figure). This means the person's least preferred cake is still preferable to any death. It is thus sometimes possible to reach a conclusion or decision without a detailed model. In such cases, gathering additional information, such as the type of the cake, will not change the conclusion and therefore provides no direct benefit, but may incur additional costs. Since the benefit does not exceed the cost, there is no *value* in acquiring extra information about the cake[1].

In general, it is infeasible to represent exactly all of the available information because this would require a model that matches the current state of the world. However, all models are by definition abstractions of reality. G. E. P. Box (1979) is widely cited as being the first to state this as, "All models are wrong; some are useful." In most cases, in order to conform the current state of information to the restrictions of a particular model, the DM must both make additional assumptions and discard available information. Consequently, there is no *absolute* truth as to what the best model is, but rather choice of a model is *relative* and depends on the available information and the decision at hand. For expensive, high-risk projects such as aerospace design, a very complicated model of uncertainty might provide the most value. For simple, low-risk project such as plastic utensil design, a very simple model of uncertainty might provide the most value.

In the context of uncertainty modeling, the goal of this dissertation is to explore, present, and develop a method for representing and managing uncertainty that is more valuable, at least in some contexts, than the methods currently used in engineering practice. These methods will not always provide greater value, but it will be shown that there are definite circumstances in which they do (0). It will also be shown that the methods have specific benefits over traditional methods that readily generalize to a broad

---

[1] It is of course possible that the intervals of preference for cake and death overlap. One example is the comparison of a cake containing a poison that leads to an agonizing death compared to a death in which the DM dies peacefully in his or her sleep. A second example is a suicidal DM, who would prefer death over any other earthly indulgences, including cake.

class of problems (Chapter 6), although the specific tradeoffs between cost and benefit are problem specific and cannot be generalized.

In this section, the notion of an uncertainty model was introduced, and the concept of model selection based on information economics was developed. One of the goals of this thesis is to identify how uncertainty should be managed in engineering design. As the next section describes, the research area of engineering design is in some ways very different from other engineering research domains.

1.4 The context of engineering design research

Much engineering research is scientific in nature, meaning that research questions are posed and refutable hypotheses are advanced. However, the *engineering design* domain is different in nature from a physics-based field such as acoustics or fluid dynamics; in practice, it is often closer to a social science than a physical science, and as such requires a different approach.

Herbert Simon makes a distinction between "natural science" and "science of the artificial" (Simon 1982). Natural sciences are descriptive, concerned with how things are, whereas the science of the artificial is normative, concerned with how things ought to be. Simon defines the artificial as things that are the result of human actions—specifically from the act of making, which he calls synthesis. Clearly engineering design is intimately connected with synthesis; design is a process that includes the act of making (synthesis) as one of its steps, but it also includes the act of observing (analysis). The aspect of Simon's theory that is important for this dissertation is the emphasis it places on the human aspect of engineering design. While engineering design includes analysis steps and has as its main goal synthesis of a product, the engineering design process is ultimately a process carried out by humans and for humans, as the resultant artifact generally provides some function to humans. Some authors acknowledge a specific subjective, rather than objective role in validation of design methods (Pedersen, et al. 2000).

According to Karl Popper (1959a), a *scientific hypothesis* must be phrased such that it can be refuted by some conceivable event; any hypothesis or theory that cannot be refuted by any conceivable event is *non-scientific*. A scientific hypothesis can be refuted by a single exception. However, humans are a crucial component of engineering design. Consequently, design is characterized by human characteristics, including beliefs, biases, values, and a strong propensity for making mistakes. Human behavior does not universally follow a clear set of rules, so there are exceptions to nearly all descriptive models of human behavior. Because a model is an approximation of reality, such exceptions should be expected; if there were no exceptions, it would be reality, rather than approximately modeling reality.

The existence of an exception to a particular model essentially refutes that model as an acceptable scientific hypothesis or theory of reality, but the model may nevertheless be useful in describing human behavior. It is for these reasons that models (rather than scientific theories) of behavior, uncertainty, and rationality are considered in this dissertation.

The process of establishing a level of trust or confidence in a model is known as model validation (Schlesinger, et al. 1979, Balci 1995, Malak and Paredis 2004). Usually some models are better than others for a particular application, even though they are all known to be imperfect. It is therefore important to identify in what ways and under what conditions one model is superior to another. Often this is a practical consideration—which model yields the most useful results for the scenario under consideration. At the same time, a valid model should have some theoretical grounding within a particular context or set of assumptions.

Research in the social and cognitive sciences has presented normative models—or normative theories—of human behavior. These models essentially state how a rational individual should act under specific circumstances. Researchers typically construct these models by beginning with a particular set of axioms and then developing a self-consistent theory from those axioms. However, there is not always universal agreement about the initial axioms. The theoretical "correctness" of these axioms is often philosophical or even theological in nature, and as such, no universal "correctness" can be proven or refuted; only intuitive and consistent arguments can be constructed and then the practical value of employing theories based on them can be demonstrated.

The practical value of the methods depends on the decision maker's preference and is a subjective, problem-dependent attribute of the method. As such, there is no point in attempting to employ purely objective, scientific tests for or against a particular method. It is necessary to combine practical and theoretical arguments, such as practically arguing for a set of assumptions, demonstrating in a scientific manner the consequences of these assumptions compared to other assumptions, and then practically comparing the consequences and their further implications.

In light of these observations, the questions posed in this dissertation are referred to as motivating questions, and proposed answers are advanced rather than pure scientific hypotheses that are directly refutable by objective observation. In the next section, the specific motivating and research questions that guide this research are developed.

1.5 Motivating questions

It was explained in the previous section that engineering design research is different from research into the natural sciences. It is more of a hybrid, combining aspects of natural sciences, social sciences, and the science of the artificial. It is thus appropriate to formulate motivating research questions and answers. These motivating questions can still be answered in a systematic and rigorous manner, and in some cases a purely scientific approach can be adopted. In other cases, more philosophical, intuitive, subjective, or practical arguments must be made.

The introduction to this chapter presents the main motivating question of the dissertation:

How should engineering designers manage information to support decision making under uncertainty?

This question is now decomposed into five secondary motivating questions. The overall essence of the questions and their connection to the primary motivating question and to each other is loosely illustrated in the followed paragraph. The specific questions and hypotheses are introduced in the following sub-sections.

As discussed in Section 1.3, the representation of uncertainty is a modeling process—a decision maker seeks to build a model of uncertainty that is appropriate for his or her specific problem. The basic story of the dissertation centers on the development of this model and proceeds in the following manner:

1. When constructing a model, it is first necessary to understand what one is trying to model, in this case asking what are the fundamental characteristics of uncertainty in the context of engineering design?
2. Once the nature of uncertainty is described, what is the most appropriate and general model of uncertainty for engineering design?
3. Given several potential models from which to choose, how should engineers compare alternative models of uncertainty?
4. Once a model is chosen, how can this model be used to support the decisions?
5. Finally, given an existing state of knowledge, a model of uncertainty associated with it, and an approach for solving the original decision problem, *how should one decide whether to proceed with the problem formulation phase and collect more information or whether to proceed to problem solution phase*?

These five questions are adapted into specific motivating questions in the following sections. The motivating questions, answers, validation, and structure of the dissertation are summarized in Figure 1.5 on the next page.

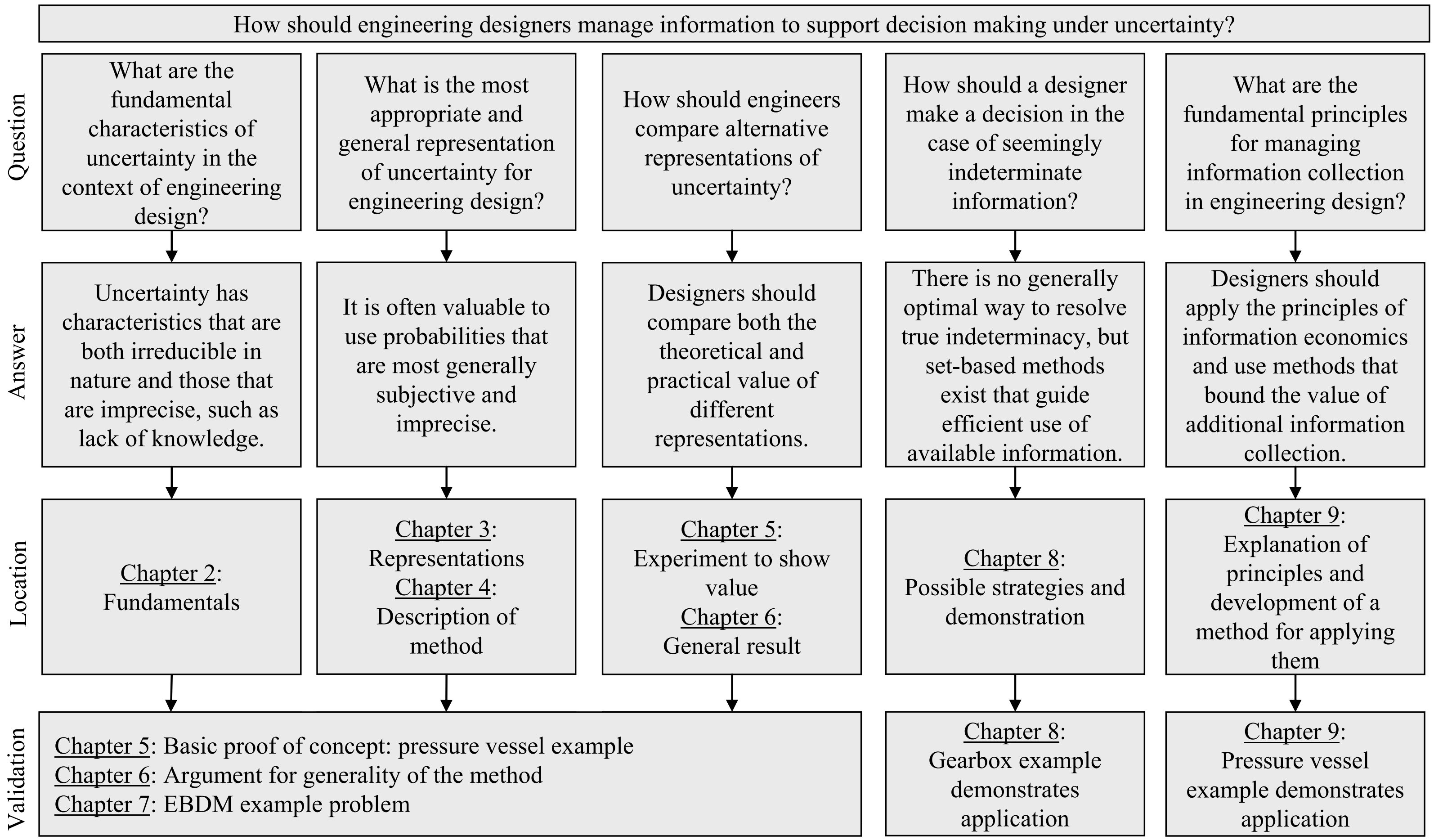


Figure 1.5. Research overview

1.5.1 Motivating Question 1

Question (Q1):	What are the fundamental characteristics of uncertainty in the context of engineering design?
Answer (A1):	Uncertainty can have characteristics of being irreducible (or random) in nature as well as characteristics that are related to a reducible lack of knowledge (or imprecision) in available knowledge.

This philosophical question is addressed directly in Chapter 2 of the dissertation. The general argument is that there are two important characteristics of uncertainty in engineering design—one characteristic is defined as irreducible uncertainty and the other is defined as imprecision, as explained in Chapter 2.

The arguments in Chapter 2 are mostly intuitive in nature. They appeal to human experience with uncertainty and draw from existing literature. More extensive practical validation of this hypothesis is achieved in conjunction with a chosen representation of uncertainty, the subject of Motivating Question 2.

1.5.2 Motivating Question 2

Question (Q2):	What is the most appropriate and general model of uncertainty for engineering design?
Answer (A2):	It is often valuable to use probabilities that are most generally subjective and imprecise.

Many models of uncertainty have been proposed. Chapter 3 contains a review of several models and a discussion of their applicability to engineering design problems. The conclusion is that the potential value of using probabilities that are subjective and imprecise in nature is high. In Chapter 4, the motivation for, the meaning of, and mathematics of a particular formalism (called probability bounds analysis) for representing and computing with uncertainty are introduced. This formalism, a refinement of the imprecise probability model of uncertainty described in Chapter 3, is adopted as the primary representation of uncertainty in the thesis.

Together, the answers (A1 and A2) to the first two motivating questions represent the core contributions of the dissertation, while the remaining questions and hypotheses are necessary and logical extensions of this core. The validation of A2 is coupled tightly with the validation of A1, since the definition of uncertainty and the model of uncertainty are inherently linked. For example, a model could capture certain characteristics of uncertainty very well, but if it neglects other important aspects, it is not a complete model. Conversely, a model that assumes characteristics that do not exist is equally troublesome.

As such, the validation of A1 and A2 are considered together. The appropriate means for comparing models of uncertainty is the subject of Motivating Question 3. The validation of A1 and A2 is presented through an example proof of concept (Chapter 5, which uses the method developed in A3), a general result (Chapter 6), and an example of the general result (Chapter 7).

1.5.3 Motivating Question 3

Question (Q3):	How should engineers compare alternative models of uncertainty?
Answer (A3):	Designers should compare both the theoretical and practical validity of different models.

Multiple models of uncertainty have been proposed in the literature. In order to choose an appropriate model of engineering design, it is necessary to compare various models to each other. The comparison should be performed from both the practical and theoretical perspectives adopted in this thesis.

The practical validity of a model in a particular engineering design problem is reflected in the practical results he or she produces. If an engineer produces a bad design, it does not matter how theoretically valid his or her approach to design was; the design is still a failure. The slight caveat is that uncertainty can cloud things. A bad outcome does not imply a bad decision; it may just reflect bad luck. When choosing a model, an engineer can ask things such as, can the use of one model lead to a better design than the use of another model can, or is one model so complex and costly to use that any benefits it provides are outweighed by the high costs? A method for making such comparisons is developed in Chapter 5. Chapter 5 also contains a comparison between the imprecise probabilities approach and a traditional, maximum likelihood estimate, precise probability based approach that serves as validation of A1 and A2.

In addition to practical comparison, a theoretical comparison should be performed in order to assess the general applicability of the model to a larger class of problems. More fundamentally, an analyst should have a true understanding of what is being modeled in order to choose a good model (Parry 1996). For example, which model more appropriately captures the true characteristics of uncertainty? Often this question will have an ambiguous answer, as most models have advantages and disadvantages that can be difficult to trade-off. However, in some cases, one model will be more general than another model, meaning its contains the other model as a special case but also can capture other relevant aspects of the uncertainty. In this situation, the more general model has a definite theoretical advantage over the less general model. The theoretical advantage of PBA over traditional decision analysis is discussed in Chapter 6 and Chapter 7.

The validation of A3 is performed from an intuitive and economic perspective. The hypothesis itself is not very controversial; however, the hypothesis is rarely put into practice. This dissertation essentially shows that both practical (Chapter 5 and Chapter 7) and theoretical (Chapter 6 and Chapter 7) comparisons of uncertainty models are possible. These comparisons are then used to justify the value of considering an imprecise probability model in engineering design.

1.5.4 Motivating Question 4

Question (Q4):	How should a designer make a decision in the case of seemingly indeterminate information?
Answer (A4):	There is no generally optimal way to resolve inherent indeterminacy, but set-based decision rules exist that guide the efficient use of available information.

A particular model of uncertainty has no practical value to a designer if there is no way to use the model to support decision making. It will be shown in Section 4.6 that the use of imprecise probabilities can yield intervals of expected utility for decision alternatives. If these intervals overlap, there is apparent (and sometimes actual) indeterminacy in the optimal action. Chapter 8 contains a discussion and demonstration of set-based methods for managing decision alternatives that guide efficient use of available information and reduce the apparent indeterminacy in the problem.

Validation of A4 is achieved in Chapter 8 via theoretical arguments, literature references to other domains, small example scenarios, and the example of the design of a gearbox.

1.5.5 Motivating Question 5

Question (Q5):	What are the fundamental principles for managing information collection in engineering design?
Answer (A5):	Designers should apply the principles of information economics and can use methods that bound the value of information collection.

Once probabilities are allowed to be subjective and imprecise, an obvious question arises: how imprecise can probabilities be and still allow for good decision making? This is a sub-question to a more general question: what are the fundamental principles for managing information collect in engineering design?

The collection of information requires the expenditure of resources, whether this involves performing more experiments, building more prototypes, or creating additional simulations. At the same time, additional information can lead to better decisions, thereby increasing the payoff of a design. The fundamental principle of information economics is that information should only be collected if its expected benefit exceeds the expected costs. This is a fundamental and widely accepted principle, but existing methods for assessing the value of information rely on assumptions that are often not valid in engineering design. Chapter 9 contains a description of these limitations and the development of a method for bounding the value of future statistical data collection.

Validation of A5 is achieved in Chapter 9 through theoretical arguments and an example problem that relates the results from a hypothetical experiment with realistic information constraints to a hypothetical experiment with complete information. The basic conclusion is that the method and principles can guide decision making, but without arbitrary resolution of indeterminacy, the methods do not provide a complete means to make decisions on information collection. However, the contributions of Chapter 9 establish a strong foundation for future work in the area of managing information collection in engineering design, work that is being pursued concurrently by other researchers (Ling 2006).

1.6 Organization of the dissertation

The relationship between the chapters of the thesis and the motivating questions and hypotheses was described throughout Section 1.5 and summarized in Figure 1.5. Chapters 2-7 form the core of the dissertation. In these, the nature of uncertainty is addressed (Chapter 2), and various models of uncertainty are considered (0). A particular model (developed by other researchers and called probability bounds analysis) is described in Chapter 4. In Chapter 5, the design of a pressure vessel using this model is compared to a traditional, best-fit precise probability model. The ability of probability bounds analysis to serve as a global sensitivity analysis is developed theoretically in

Chapter 6, and then the process of using probability bounds analysis in a design problem, including as a sensitivity analysis, is demonstrated in Chapter 7.

Chapter 8 and Chapter 9 are logical extensions to the core of the dissertation. The issue of decision making is addressed and demonstrated in Chapter 8. This topic is crucial for widespread application of the methods discussed in the rest of the dissertation. The topic of Chapter 9 is information economics and the collection of information to support design decisions. This chapter servers as a gateway looking forward. The adoption of imprecise probabilities allows for a new method for bounding the value of future statistical data collection to be developed (presented in Chapter 9), but many issues area left unresolved. As such, the chapter points towards a new direction for future work, demonstrates an additional use of imprecise probabilities and probability bounds analysis, and presents preliminary results in this new direction.

CHAPTER 2:

UNCERTAINTY IN ENGINEERING DESIGN

This subject of this chapter is the first motivating question of the dissertation:

What are the fundamental characteristics of uncertainty in the context of engineering design?

In Section 2.1, uncertainty is defined. In Section 2.2, the notion of different states of uncertainty is introduced. Section 2.3 contains a discussion of types of uncertainty and characteristics of uncertainty. Finally, sources of a particular characteristic of uncertainty, called imprecision, are introduced in Section 2.4

2.1 Definition of uncertainty

In this dissertation, uncertainty is viewed in the context of decision theory and is defined, following Nikolaidis (2005), indirectly from the definition of certainty. Nikolaidis defines certainty as the condition of knowing everything necessary to choose the course of action whose outcome is most preferred. A decision-maker's uncertainty is defined as the gap between certainty and the decision-maker's present state of information—the information the decision-maker currently has available for decision making, which is a slight refinement of Nikolaidis's definition[2]. As an example, consider

---

[2] The relative sizes of components in Figure 2.1 have no intended meaning; the purpose of the figure is only to show how the components are related to each other.

the truth that $X = 12$. If the DM knows this, then the DM is in a state of certainty. If the DM only knows that $10 \le X \le 14$, then there is uncertainty in the true value of X.

2.2 Recognition of different states of uncertainty

Decision theory has long differentiated between decision making with known probabilities[3] and decision making without knowledge of probabilities[4] (Knight 1921). Researchers also have explored the middle ground of partial or incomplete knowledge of probabilities (Cannon and Kmietowicz 1974), such as ordered probabilities (Fishburn 1964) and linear constraints on the probabilities (Kmietowicz and Pearman 1984). Other literature has examined incomplete or partial information (see (Weber 1987) for a review) in the context of imprecisely characterized preferences (Otto and Antonsson 1992, Carnahan, et al. 1994, Seidenfeld, et al. 1995) and unknown weights in multi-attribute decision making (Kirkwood and Sarin 1985). While many of these distinctions are made under the assumption of probability as the model of uncertainty, they also indicate that there is more to uncertainty than just probability.

This contradicts a common assumption [see for example (Lindley 1982a, Winkler 1996)] that traditional probability theory is the only acceptable language and mathematics of uncertainty. In the standard terminology of decision theory (Knight 1921), the word uncertainty explicitly means without knowledge of probabilities. If probabilities cannot capture all uncertainties, *what is uncertainty* (motivating question 1), and *how can uncertainty be represented* (motivating question 2)? These questions are the topics of this chapter and the following, respectively.

---

[3] What Knight calls *decision making under risk*. This definition of risk is not used in this dissertation.

[4] What Knight calls *decision making under uncertainty*. This definition of uncertainty is not used in this dissertation.

2.3 Types of uncertainty

Many authors attempt to subdivide uncertainty into specific types (Der Kiureghian 1989, Casti 1990, Helton 1994, Hoffman and Hammonds 1994, Rowe 1994, Ferson and Ginzburg 1996, Hofer 1996, Hora 1996, Parry 1996, Pate-Cornell 1996, Cullen and Frey 1999, Oberkampf, et al. 2001, Oberkampf, et al. 2002b, Dai, et al. 2003, Haukaas 2003, Nikolaidis 2005). The motivation for such distinctions is that intuitively it seems that there are things that are inherently unknowable until they are realized (e.g., the outcome of a basketball game that takes place tomorrow), and there are also things that are in theory knowable but may be unknown to the decision maker (e.g., the cumulative stats for both teams in the basketball game as of now, or the outcome of the game the last time the teams played). In this section, types of uncertainty are presented and a distinction is made between imprecision and irreducible uncertainty.

2.3.1 Reducible and irreducible uncertainty

The uncertainty associated with unknowable things is often referred to as irreducible uncertainty (Der Kiureghian 1989). The uncertainty associated with things that are knowable but are currently unknown is often referred to as reducible uncertainty. Philosophically, the distinction between the two is troublesome. For example, is the uncertainty about tomorrow's weather reducible or irreducible? From one perspective, the weather cannot be known until it happens. On the other hand, it might be possible to predict tomorrow's weather exactly using thermodynamics and hydrodynamics—if the current state of every atom on Earth and all physical laws were known exactly, and if tomorrow's weather is a deterministic function of today's weather, then tomorrow's weather is knowable. From yet a different perspective, perhaps quantum effects or Heisenberg's uncertainty principle are factors. These possibilities lead again to the question, is accurate weather prediction inherently *impossible* or just presently *impractical*? There is not necessarily a clear answer.

2.3.2 Aleatory and epistemic uncertainty

Another distinction frequently made in the literature is between aleatory uncertainty and epistemic uncertainty (Parry 1996, Oberkampf, et al. 2002b, Dai, et al. 2003, Haukaas 2003). The term *aleatory* uncertainty comes from the Latin *aleator* for dice thrower or gambler, and refers to uncertainty that arises from a random process. Other authors refer to the same concept as inherent variability, or just variability. It is almost universally accepted that aleatory uncertainty is best modeled using probabilities. It is recommended that readers who are not familiar with both the frequentist and subjective interpretations of probability refer ahead to Section 3.3.3, as a basic understanding of these concepts will be useful throughout the dissertation.

The counterpart to aleatory uncertainty is epistemic uncertainty, with the term *epistemic* arising from the Greek *episteme* for *knowledge*, which is also linked to the verb *to understand*. Epistemic uncertainty is caused by a lack of knowledge; this lack of knowledge is a state of the analyst or decision-maker, rather than a state of the physical system under consideration. The distinction is motivated by the “location” of the uncertainty—in the decision-maker or in the physical system.

Many authors align epistemic uncertainty with reducible uncertainty and aleatory uncertainty with irreducible uncertainty, but the mapping is not exact. If one accepts the existence of aleatory uncertainty, it is certainly a form of irreducible uncertainty. However, the outcome of future deterministic event could be uncertain, and this uncertainty could be irreducible without being an inherently random process. For example, the winner of the best picture award at next year’s Academy Awards is clearly unknown, but it is almost certainly not an inherently random process. Similarly and more relevant to engineering design, a future detail design decision that affects the total payoff of the design is uncertain during conceptual design, yet (assuming a systematic design process) that detail design decision is not random but rather guided by rigorous analysis and the input of expert engineers.

As with irreducible and reducible uncertainty, the difference between aleatory and epistemic uncertainty is not always clear. The dimensional errors in manufacturing processes are often modeled as the result of a random process, and therefore as an aleatory uncertainty. However, in many cases it can be argued that the manufacturing process is a deterministic, though perhaps chaotic system that obeys the deterministic rules of Newtonian physics. Rather than the process being random, it is more likely the input (environmental) parameters to the process that lead to the appearance of random behavior. The difficulty in determining whether an uncertainty is aleatory or epistemic is another argument against the actual existence of different uncertainties, but it does not imply that such distinctions have no value.

Despite philosophical questions about the existence of aleatory uncertainty, it is clear that engineers frequently model processes as inherently random. As discussed in Section 1.3, models of uncertainty, like all models, will not be exact replications of reality. It is thus necessary to consider aleatory uncertainty in the practical context of modeling (Parry 1996, Winkler 1996).

2.3.3 Practical usefulness of distinguishing different characteristics of uncertainty

Despite the philosophical arguments against inherently different types of uncertainty, many modern authors, such as (Ferson and Ginzburg 1996, Hofer 1996, Parry 1996) and including skeptics of a fundamental distinction between types of uncertainty such as (Winkler 1996), agree that it is useful in practice to make such distinctions when possible . This is especially true when considering building models of uncertainty for use in engineering decision making, because there are some characteristics of uncertainty that are starkly different and should be modeled as such.

For example, the strength of different samples of a particular material differ due to small fluctuations in the manufacturing process. Engineers frequently choose to model the strength of a novel material as a random variable. Assume an engineer knows from historical data that most materials are well modeled by a normal distribution with truncated tails (a truth assumed in this dissertation for illustration only and with no implied relevance to material science). It is also assumed that the true process is stationary. Consequently, the engineer chooses to model the strength of this novel material as a normal distribution.

Based on the engineer's modeling decision, there is clearly some type of irreducible uncertainty about the strength of a particular material sample; the engineer cannot known the specific material strength of a future sample until that material sample is manufactured. However, there is also a type of reducible uncertainty present because the engineer does not know the parameters (mean and standard deviation) that describe the particular normally distributed stochastic process . This uncertainty can be reduced by performing an experiment, such as creating 100 samples of the new material and measuring the strength of each and then using the sample mean and standard deviation to estimate the overall population parameters.

Extending this, 100 samples is not guaranteed to lead to an accurate characterization of the mean and variance. Collecting 1000 samples would be better, and 10,000 even better still. This additional information collection is not reducing the uncertainty from the random manufacturing process; it is merely reducing the uncertainty in how well the random process is known to the engineer. If the engineer can afford to perform 1 million tests, he or she will know the parameters of the distribution even better. In practice, collecting samples costs resources, and resources are limited. Consequently, an engineer generally cannot get arbitrarily close to the true parameters of a random process. It is therefore necessary to make tradeoffs between the cost of acquiring additional information and the benefits of such information. In order to guide such tradeoffs, it is necessary to express the lack of knowledge about the true parameters—uncertainty that will be called imprecision.

2.3.4 Imprecision and irreducible uncertainty

In this dissertation, the definitions summarized in Figure 2.1 are adopted. Specifically, the definition is adopted that irreducible uncertainty accounts for the gap, shown in Figure 2.1(b), between certainty and a state of precise information, defined as the state of having acquired all information about a particular model of irreducible uncertainty available at any price. As described in the previous section, even if a quantity is assumed to be random, the type of the distribution (e.g. normal) and its parameters (e.g. mean and variance) still need to be determined. If the distribution type and parameters are known perfectly, then the irreducible uncertainty is known precisely.

The gap between the present state of information and a state of precise information about the irreducible uncertainty, shown in Figure 2.1(b), is defined as imprecision. The terms aleatory and epistemic are specifically avoided in part because they have been used primarily in an attempt to differentiate the inherent nature of different uncertainties, rather than focusing on how human decision-makers should manage and model uncertainty in engineering design. Additionally, there are irreducible uncertainties that are not inherently random in nature, such as the examples noted earlier of future sporting event, future Academy Awards, and future design decisions.

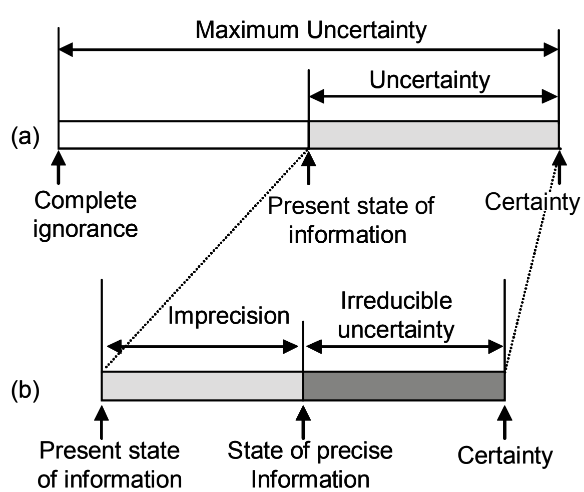


Figure 2.1: Characteristics of uncertainty, adapted from (Nikolaidis 2005)

The term *imprecision* is chosen over *reducible uncertainty* because it provides additional clarity in meaning. For example, it is useful to discuss the difference between something being known *precisely* and *imprecisely*, terminology that is not as obviously aligned with the word reducible. *Precise* is defined as *exactly or sharply defined or stated* and precision is the quality or state of being precise (Merriam-Webster 1993). Imprecise is simply not precise and by extension, *imprecision* is *the quality or state not being exactly or sharply defined or stated*. This is exactly what needs to be captured in engineering design. If an engineer lacks knowledge about some parameter, it is not sharply defined; as more information becomes available, the parameter becomes more sharply defined until it is known precisely. Until that point, the presence of imprecision represents the lack of complete information, which is also an opportunity to collect more information; when an engineer fails to recognize imprecision, he or she may also fail to recognize the need for additional information.

As with the distinctions of aleatory versus epistemic and reducible versus irreducible, there are cases when the distinction between imprecision and irreducible uncertainty is not clear. This does not, however, imply that they are identical, and certainly is not proof that they should be modeled in exactly the same way. However, it does suggest that as imprecision reduces, the model of total uncertainty should approach a model of pure irreducible uncertainty. The modeling of uncertainty is addressed at length Chapter 3.

Previous work has examined imprecision in preferences (Otto and Antonsson 1992, Antonsson and Otto 1995), but in that literature the word imprecision is used in a slightly different, though related way that is applicable only to preferences. This dissertation focuses on a more fundamental and general problem that subsumes imprecision in preferences, and it is assumed that preferences are known precisely.

Additional arguments for the explicit consideration of imprecision separately from irreducible uncertainty are made in the following chapter, focusing on the context of specific uncertainty representations (such as probability theory). However, before discussing representations, the consideration of imprecision is justified by describing the many sources of imprecision in engineering design.

2.4 Sources of imprecision in engineering design

Almost every aspect of the engineering design problem introduces imprecision[5]. Figure 1.3 on page 8 provides a nice context for exploring these sources. Specifically:

1. Sequential decision making introduces imprecision because the results of future decisions are unknown.
2. Statistical data from finite samples of environmental factors are inherently imprecise.
3. Bounded rationality leads to imprecise subjective probabilities.
4. Expert opinion and judgments are not precise, due to lack of information or conflict.
5. Behavioral simulations and analysis models are imprecise abstractions of reality.
6. Preferences may be imprecise due to bounded rationality or non-stationarity.
7. Numerical implementation of these models introduces additional imprecision.

The following sub-sections are an elaboration on how these sources introduce imprecision into the design process.

[5] Most of this section was previously published in a workshop paper and presentation (Aughenbaugh and Paredis 2006b)

2.4.1 Sequential decision making

As noted in Section 1.1.2, the complexity of the design problem makes it impossible to arrive at an optimal design in one step. Instead, the process is divided into a sequence of decisions. This process is illustrated using a simple design problem with two design variables: vehicle type and engine type. There are two options for vehicle type: car or bike. There are three options for engine type: gasoline engine, diesel engine, or electric motor. If the DM chooses the design in one step, he or she would choose from the set of six *design alternatives* shown in Figure 2.2. In the context of this example, each of these design alternatives is a fully detailed design of a final product.

In order to choose the best design out of these six, the DM would need to evaluate and compare all six. While easy in this simple example, it is impractical to enumerate and evaluate all design alternatives by considering all possible combinations of all solution principles for all the subsystems of a complex product. Consequently, the decisions are broken down into sequences to allow for efficient exploration of the design space. For example, in the previous vehicle design example, a DM can follow a sequential approach in which he or she first chooses the vehicle type, and then the engine

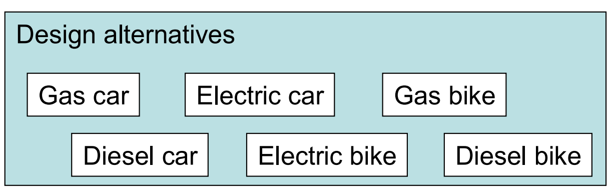


Figure 2.2. One stage decision

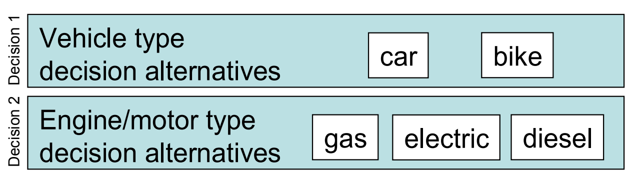


Figure 2.3. Sequential decisions

type, as shown in Figure 2.3.

Note that it is important here to distinguish clearly between *decision* alternatives and *design* alternatives. A design alternative is one of the possible complete product design specifications (recall Figure 2.2), while each decision alternative is a specific option for a specific decision and corresponds to a set of design alternatives. For example, when choosing the vehicle type, the DM has two *decision alternatives*: car or bike. Each of these decision alternatives actually corresponds to a *set* of *design alternatives*, as shown in Figure 2.4.

The choice of decision alternative *car for vehicle type* includes the gas car, diesel car, and electric car design alternatives, because the decision about vehicle type will be followed by the decision about engine type. Once a decision is made to pursue, for example, a car design rather than a bike, the DM does not need to consider explicitly the design alternatives of gas bike, electric bike, and diesel bike; these design alternatives are *eliminated* from consideration.

One limitation of a sequential decision process is that decisions often are coupled. In general, one really needs to know the outcome of future decisions to select the best (or most preferred) decision alternative for the current decision. For example, a fully designed car will have a certain maximum horsepower, but this certain value is unknown when the vehicle type decision is made (Figure 2.3), because it depends on the future

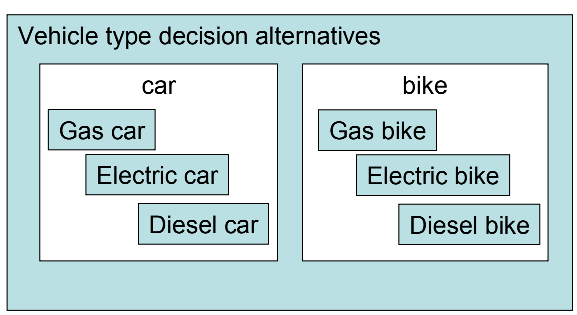


Figure 2.4. Decision alternatives and sets of design alternatives

design decision of engine type. The set of car designs in Figure 2.4 has multiple horsepower maxima, each corresponding to a sub-design (gas car, electric car, and diesel car). Thus, when selecting type *car rather than bike*, a DM is not selecting a precisely characterized horsepower, but rather a set or interval of horsepower. In a more complex problem, imprecision will remain once the engine type is chosen because a particular engine type is a set of designs. For example, even if a gas engine is chosen, characteristics such as horsepower, torque, mass, and fuel efficiency will be inherently imprecise because they depend on additional details of the design.

The inherent existence of sets in sequential decision making demonstrates the need to compute with intervals, sets, or information that otherwise is characterized imprecisely. However, other sources of imprecision are independent of the existence of sets of design alternatives. These may have different characteristics and may affect the design process differently, as described in the following.

2.4.2 Limited statistical data

In order to support decisions, engineers frequently gather statistical data about uncontrollable factors such as the environment. Such quantitative data gives an illusion of being well-characterized, but actually it is inherently imprecise. Assume one needs to design a pressure vessel, and the vessel will be made of a new type of steel for which the yield strength X is not well characterized. Engineers have strong theoretical evidence that the material strength is normally distributed, but they do not know the mean μ or variance σ^2 of the distribution. Because the material is new and testing is relatively expensive, the engineers have only measured the yield strength in a set Σ of n independent tension tests, where n is a relatively small number due a high cost of testing. These tests can at best give an estimate of the true distribution, so in addition to inherent randomness (irreducible uncertainty), engineers also face imprecision—they cannot characterize the parameters of the random variable precisely.

For example, assume the engineers have a set of 30 material strength measurements. They could use the 30 samples to estimate the true mean and variance of the distribution using standard statistics. However, these estimates ($\hat{\mu}$ and $\hat{\sigma}^2$) are exactly that—estimates. The resulting distribution $X \sim N(\hat{\mu},\hat{\sigma}^2)$ in general is not the true distribution. Alternatively, confidence intervals can be constructed on the true mean and variance at the α confidence level as follows, where n is the number of samples and s is the sample standard deviation (Hines, et al. 2003), $t_{\alpha/2,n-1}$ is the t-statistic and $\chi^2_{\alpha/2,n-1}$ the Chi-squared statistic (tables of these statistics are found in most statistics books, such as (Devore 1995, Hines, et al. 2003):

$$\left[\underline{\mu},\overline{\mu}\right]=\left[\hat{\mu}-t_{\alpha/2,n-1}\frac{s}{\sqrt{n}},\ \hat{\mu}+t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right] \tag{2.1}$$

$$\left[\underline{\sigma^2},\overline{\sigma^2}\right]=\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}},\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right]. \tag{2.2}$$

These confidence intervals for the parameters represent imprecision in the knowledge of their true values. Returning to the definition of imprecision, the parameters are not defined exactly or sharply, but rather imprecisely using intervals. With finite data (a definite practical constraint), theses parameters can never be known exactly, although an engineer may decide to model them as such. Common practice frequently does just this, assuming away imprecision; this dissertation shows that this assumption can sometimes be costly (0).

In this section, the focus was on statistical data, emphasizing a rather frequentist interpretation of probability—an interpretation in which a probability represents a long-run relative frequency. Often, a subjective interpretation is more applicable. The specific meaning of frequentist probabilities and subjective probabilities, as well as other interpretations of probabilities, is discussed in much greater depth in Section 3.3.3. Section 3.3.3 also contain a discussion of the appropriateness of various interpretations.

For now, it suffices to state that if either interpretation is adopted, there can be imprecision in the knowledge of the probabilities.

2.4.3 Imprecise subjective probabilities

Proponents of a subjective interpretation of probability assert that there is no such thing as a true or objective probability, but rather probabilities are an expression of belief based on an individual's willingness to bet (de Finetti 1974, Lindley 1982b, Winkler 1996) (see also Section 3.3.3.3). One of the subjectivists' primary arguments against a frequentist perspective is the absence of truly repeatable events, especially in practical problems. For example, the probability that Team A beats Team B in a basketball game has no real meaning under a frequentist interpretation, because that event—that particular game—will occur exactly once; there is no long-run relative frequency. In this context, the notion of a long term frequency, and even random events, is meaningless (de Finetti 1974). However, many people are willing to express their belief of who will win in terms of bets. When framed appropriately, such bets can be taken as subjective probabilities.

In order for subjective probabilities to be precise, the decision maker (DM) must fully elicit his or her betting behavior. In other words, for any given gamble, the DM must be able to state his or her fair price—the price at which the DM is willing to take either side (selling or buying) of the bet. For example, consider the toss of thumbtack (Walley 1991). A DM is asked to state the probability that the tack lands pin-up. This may require the DM to think very hard about the gamble. The DM may also want to collect additional information about the problem before making a decision. These steps cost time and other resources. However, the DM may easily be able to state the probability of the tack landing pin-up is between 0.2 and 0.8. This interval is a form of an imprecise probability, an uncertainty model introduced in Section 3.4.

2.4.4 Expert opinion

A significant source of information in engineering design are experts who use their knowledge and experience to form judgments, beliefs, and estimates (Cooke 1991, Ayyub 2001, Coolen 2004). Information from expert opinions is inherently imprecise. First, opinions may not always be cited precisely, especially when expressed in vague linguistic terms, such as unlikely, large, or poor, a case in which fuzzy set theory may have a role (Zadeh 1965, Ayyub 2001) (see also Section 3.2.1). Because an opinion about the world is not necessarily the truth of the world, opinions also can differ from person to person. Often, these opinions will conflict. For example, two experts are asked the probability that a quantity X is below 5; that is, $P\{X < 5\}$. The first expert says that $P\{X < 5\} = 0.3$ (and consequently $P\{X \geq 5\} = 0.7$). The second expert states that $P\{X < 5\} = 0.6$ (and consequently $P\{X \geq 5\} = 0.4$). Does $P\{X < 5\} = 0.3$ or does $P\{X < 5\} = 0.6$? Maybe the best characterization of the expert knowledge is to say that $0.3 \leq P\{X < 5\} \leq 0.6$, or it is imprecisely known that $P\{X < 5\} = [0.3, 0.6]$. The combination of such evidence, especially when conflicting, is an important research area, often focused on Evidence Theory (Dempster 1967, Shafer 1976, Yager, et al. 1994b, Oberkampf and Helton 2002, Mourelatos and Zhou 2005a) [see also Section 3.2.3]. This issue arises for all information theories and uncertainty models.

2.4.5 Imprecise analysis models

An important step in decision making and design is to determine the DM's preferences over design alternatives. The DM's preferences are determined by attributes (e.g. performance) of the design alternatives. As illustrated back in Figure 1.3, the determination of the DM's preferences between design alternatives involves the application of multiple models: simulation models that predict the performance of the alternatives, models for the uncertain inputs to these behavioral models, and models of the DM's preferences.

Behavioral models predict the performance of design alternatives in terms of attributes that are important to the DM, such as physical behavior, cost, and reliability. Since these models, like all models, are only an abstraction of reality, they are imprecise (Parry 1996). Specifically, although the laws of physics are known very precisely, one often makes significant assumptions when applying the laws of physics to complex geometries, or one omits certain known—but less significant—physical phenomena from the model to reduce the complexity.

For example, a model for an internal combustion engine is often abstracted into an algebraic relationship between engine speed and torque. The detailed physical phenomena (including airflow, gas-mixture combustion, friction, and inertia) are reduced into one simple algebraic relationship. This simple relationship is an idealization that may contain a significant error—the unknown or unmodeled relationships between a variety of parameters that play a role in the engine performance, such as air density, acceleration, or engine temperature. The lack of knowledge of the influence of these parameters on engine performance results in imprecision in the model's predictions. Since there is no probability distribution associated with such modeling and systematic errors, one cannot express the likelihood of occurrence for a particular error but can at best bound the size of the error, in which case the errors should be represented in terms of interval-based uncertainty.

In addition to the imprecision in the behavioral models themselves, there is often also significant imprecision in the parameter values or inputs to these models. For example, the air resistance model of a car may include a drag coefficient, which can only be determined precisely through experimentation that is more extensive. Given the limited resources (cost, time, etc.) available for experimentation, the coefficient is only determined up to certain error bounds, which introduces additional imprecision in to the model predictions. There may also be stochastic environmental noise parameters; in addition to the inherent variability of such parameters, they will be imprecisely characterized, as described in the preceding sections for statistical data or subjective probabilities.

2.4.6 Imprecise preferences

Once the performance attributes of a particular design alternative have been determined, they are combined in a preference model to form a measure (such as expected utility) of the DM's overall preference for the specific alternative, as is illustrated in Figure 1.3. Keeney and Raiffa (1993) propose a method for developing such a preference model by eliciting preferences with respect to single attributes, expressing the preferences under uncertainty in utils, and then combining the utility functions of the multiple attributes into an overall utility function. Due to resource constraints, such a complete elicitation and precise characterization is unachievable in practice. Instead, the preference model is an imprecise abstraction based on limited preference elicitations. Other literature has examined incomplete or partial information [see (Weber 1987) for a review] in the context of imprecisely characterized preferences (Otto and Antonsson 1992, Carnahan, et al. 1994, Seidenfeld, et al. 1995) and unknown weights for tradeoffs between objectives in multi-attribute decision making (Kirkwood and Sarin 1985).

There is also evidence that people cannot express their preferences well in a rational fashion. When presented with choices between which a rational decision maker should be indifferent, even knowledgeable experts with a strong background in decision theory often judge the choices differently (Tversky and Kahneman 1974). This psychological evidence suggests that the environment and manner in which a choice is posed affects the elicited action, and thus choices are not a perfect indication of preference. It is also possible that preferences are non-stationary, meaning they vary over time. Even if they are reasonably stationary over a relevant time horizon, practical and psychological evidence strongly suggest that preferences can only be modeled imprecisely.

2.4.7 Numerical calculations

This source of imprecision is probably the most familiar in engineering design, but possibly the least significant. The calculations implemented on a digital computer are only precise up to the machine's numerical precision. In practice, modern computers have a very high precision, and this effect is generally not important, especially in comparison to the other sources of imprecision in engineering design. For example, consider the use of a model to calculate some parameter. It often does not matter whether the numerical solution of this model is within 10^{-10} or 10^{-15} of the model's "true" answer, because the model being used is already imprecise; moving to 10^{-15} accuracy just means that one would know the model's wrong answer better; it provides no further insight into the true answer for the real system.

Imprecision also can arise with the use of numerical methods, which are used to approximate analytical solutions when analytical methods are unavailable. Some of these methods are not guaranteed to converge on the exact solution for certain problems, and thus introduce considerable uncertainty that an analyst must explore. Other methods converge on the true solution, but this convergence is not exact in most algorithms; there is usually a tolerance set in them as a stopping criterion. For example, an iterative method may terminate when the solution changes by less than some small amount over several iterations. Consequently, the solution is known imprecisely. While these computational issues are of some interest, it is again believed that the imprecision they introduce often is inconsequential compared to imprecision from other sources.

2.5 Summary

The first motivating question of the dissertation asks *what are the fundamental characteristics of uncertainty in the context of engineering design*? The recognition of a lack of information, including recognition of the opportunity to reduce that lack of information, is an important motivation for recognizing two characteristics of uncertainty. Essentially, if imprecision is ignored, then everything appears to be precise—exactly or sharply defined. If everything is always known precisely, then there is never a need to collect information. It then follows that everything is already known, a clear contradiction with reality. Consequently, the recognition of imprecision in information is a necessary step in uncertainty and information management.

The remainder of this dissertation contains a search for methods to incorporate imprecision into engineering decision making and a demonstration of the value of so doing. An important step in the process is establishing a mathematical model of uncertainty that allows distinctions between irreducible uncertainty and imprecision to be made, such that the two characteristics of uncertainty to be considered separately. This is the subject of Chapter 3.

CHAPTER 3:

MODELING UNCERTAINTY

In the previous chapter, the nature and sources of uncertainty in engineering design were considered. If engineers are to make good design decisions, they need means for representing these uncertainties and incorporating them into the decision process. Along these lines, this chapter contains a preliminary answer to the second motivating question of the dissertation:

What is the most appropriate and general model of uncertainty for engineering design?

Section 3.1 contains a discussion of the importance, in addition to clearly defined axioms and calculi, of well-defined interpretations for any mathematical representation or model. Section 3.2 contains introductions and discussions of several uncertainty model, specifically fuzzy sets, possibility theory, and evidence theory. In Section 3.3, traditional probability theory is introduced. Specific attention is paid to interpretations of probability and the applicability of these interpretations to engineering design. In Section 3.4, the theory of imprecise probabilities is described and an argument is made in favor of imprecise probabilities as a useful, general model of uncertainty. Finally, Section 3.6 contains a development of a general mathematical formulation of decision making under uncertainty.

3.1 The importance of well-defined interpretations

An important distinction to make when modeling uncertainty is between a mathematical representation of uncertainty and an interpretation of uncertainty. Klir and

Smith (2001) define four distinct levels that must be addressed for a theory of any type of uncertainty to be complete:

1. Define an appropriate mathematical representation of that uncertainty
2. Develop a calculus by which that uncertainty can be manipulated
3. Define a meaningful way of measuring the uncertainty in any situation formalizable by the theory
4. Develop methodological aspects of the theory, including procedures for making the various uncertainty principles operational within the theory

For example, consider traditional probability theory. This is a specific theory of uncertainty with clearly defined axioms (definition of representation) and a clearly defined calculus that is consistent with those axioms (see Sections 3.3.1 and 3.3.2 for more). To many scientists and engineers, the combination of these axioms and this calculus are probability; anything that satisfies these axioms and obeys this calculus can be considered a probability. But how can the probability of some event be measured? What does a particular probability mean? These questions are not answered by the axioms and calculus.

Most engineering design literature focuses on the mathematical representations and calculus of methods, with much less attention given to the interpretation of the results of these calculations. However, if one does not know what something is or how to measure it, then he or she cannot use it to help make a decision, regardless of how well the mathematics are defined.

In order for a formalism to be acceptable for practical use, it is not necessary that only one interpretation be consistent with the axioms and calculus. However, it is necessary that at least one unambiguous interpretation is consistent with the axioms and the calculus. This interpretation must be clearly defined such that when the formalism is used, the interpretation can be clearly stated and clear meaning conveyed. According to an operational perspective, a quantity is only defined unambiguously only once the process of measuring it—step 3 in Klir and Smith (2001)—is defined, as explained in the following subsection.

3.1.1 Operational definitions in science

In order for a model of uncertainty to be complete, a clearly defined *operational interpretation* what that model captures is necessary. The notion of an operational definition was first formalized by Bridgeman (1927). Bridgeman's work was primarily in the area of the philosophy and logic of "modern" physics, where "modern" at the time (ca. 1927) meant resolving the issues that existed between classical physics and the newer theories of relativity. For example, Einstein's exposition of relativity called into question traditionally held notions of length and time. Physicists were left asking, "what is length?"

Bridgeman writes the following:

> We evidently know what we mean by length if we can tell what the length of any and every object is....To find the length of an object, we have to perform certain physical operations. The concept of length is therefore fixed when the operations by which length is measured are fixed: that is, the concept of length involves as much as and nothing more than the set of operations by which length is determined. In general, we mean by any concept nothing more than a set of operations; *the concept is synonymous with the corresponding set of operations*.

Stevens (1935, 1936, 1939) [see (Hardcastle 1995) for a synopsis of this work], in writings focusing on psychology rather than physics, developed a similar notion of *operationism* that emphasized the need for agreement in science as to what something is. In this dissertation, operational is taken in the spirit of Bridgeman's definition. As such, an operational definition of a quantity is a set of operations, or a procedure, by which that quantity can be measured. If a procedure is clearly defined, then the quantity is uniquely determined by that procedure.

If a concept is defined using multiple procedures, then there is a chance of ambiguity and confusion, because different procedures may lead to different results—one procedure may measure one length, and another procedure may lead to a different length Bridgeman (1927) uses an example of measuring length of an object using a pole of set length versus measuring the object's length using light and optics; the optical method can deviate from the pole approach due to relativistic effects from the motion of Earth's rotation. When conveying a length measurement, absolute rigor requires conveying the operational definition to which the length corresponds.

3.1.2 The role of the observer in science

A clearly stated operational definition does not remove all ambiguity. For example, assume that the definition of length included an operation such as "count the number of times your left foot fits linearly along a side of the object." Although simple, this is an operational definition of length, but confusion can arise when two people attempt to measure the length of an object.

For example, Analyst 1 measures a box and determines it is a cube with sides of length 2 feet. Analyst 2 measures the same box and determines it is a cube with sides of length 2.2 feet. How can the same box have different lengths? It turns out that each analyst uses his (or her) own left foot as the base unit. Is this an operational definition? In a sense it is operational, but it is not absolute. Based on this definition, the length of an object varies based on who measures it.

The introduction of the theory of relativity into physics emphasized the role of the observer in science. This was a consequence of Einstein's analysis of the meaning of simultaneity of events (Bridgeman 1927). Bridgeman writes that Einstein showed that

"…the operations which enable two events to be described as simultaneous involve measurements on the two events made by an observer, so that 'simultaneity' is, therefore, not an absolute property of the two events and nothing else, but must also involve the relation of the events to the observer." It is thus clear that observers play a role in science, a fact that is frequently overlooked. Many researchers seek a purely objective view of science in which all observers observe the same events and reach the same conclusions. However, the role of the observer in a measurement actually does not exclude that measurement from a scientific process.

3.1.3 Standards of measurement

The example definition of length in terms of one's own left foot is operational and theoretically defensible, but practically it is not very useful due to a lack of transferability and difficulty in repetition of measurements. Analyst 1 could repeat Analyst 2's measurement of length, assuming that Analyst 1 has access to Analyst 2's measuring stick, i.e. his or her left foot. To be more practical, physical quantities usually are operationally defined in terms of standard reference materials, such as those maintained in the United States by the National Institute of Standards and Technology. These standard reference materials are certified as having specific characteristics and are used as calibration standards for other measuring devices. In essence, a standard reference material for length would define exactly what a "foot" is, and all length measurements reported in units of feet should match the prescribed operations for measuring length in term of this standard.

It is noted that a quantity could have multiple operational definitions and even multiple standards; for example, anyone can define their own operations for measuring length, but this definition will not be understood—and worse, could be misunderstood as something else—by most other people. Over time, this definition could become the new standard definition of length. For example, the definition of the unit of length called a meter has changed over time, variously being defined as the length of pendulum with period of 2 seconds, one ten-millionth of the length of the earth's meridian along one-fourth the polar circumference, a particular platinum *metre bar* placed in the National Archives in France, and the length of the path traveled by a particular wavelength of light in vacuum during a time interval of 1/299,792,458 of a second (2006). To be precise, someone must declare what standard meter he or she is using. However, only the existence of operational definitions and standards makes such clarity possible.

3.1.4 Operational definitions for uncertainty models

Cooke (2004) provides an clever example of the importance of clear definitions. He begins his paper with the following anecdote:

> I asked Didier Dubois at a 1996 meeting of the European Fusion work group in Lecoutre, France: "How many legs does a squizzel have?"
>
> He answered: "First tell me what a squizzel is."
>
> Right answer. But instead of telling him, I said: "Well, just use your own idea of what you think a squizzel is, and tell me how many legs it has."

Interpreting this example, if Cooke and Dubois have entirely different definitions of what a squizzel is, then Dubois' answer (regarding the number of legs) is useless to Cooke. Perhaps Dubois defines a squizzel as a type of spider, and Cooke defines a squizzel as a type of primate; then Dubois's belief in the number of legs a spider has is of absolutely no help to Cooke's goal of determining the number of legs a primate has. Moreover, it would be irresponsible for Cooke to use Dubois' answer to guide an important decision, because Cooke has no idea of what Dubois was thinking when he answered Cooke's question.

This example focused purely on the semantics of language, but the same holds for the semantics of formalisms: if two individuals cannot communicate a particular meaning of information, then the information is useless. Consider the following survey questions:

What is the probability that it rains tomorrow?

What is the possibility that it rains tomorrow?

What is your belief that it will rain tomorrow?

How likely is it to rain tomorrow?

Perhaps the responses given are, respectively: 0.6, 0.8, I think it will rain, and likely. What do these mean? How might they affect a decision maker's (DM's) actions?

Consider another example inspired by Cooke's squizzel example. A DM is presented with two gambles:

- Gamble 1 pays $100 if event A occurs, zero dollars otherwise.
- Gamble 2 pays $150 if event B occurs, zero dollars otherwise.

The decision maker is also told that the *fadizzle* of event A is 0.6 and the fadizzle of event B is 0.3. Which gamble should the DM choose? Clearly that depends at least in part on what a fadizzle is. If the DM, the reader, and the author all have different operations for measuring fadizzles, then a fadizzle is not a useful measurement unit or description of an event. If a clear, operational definition is developed, then the DM, reader, and the author could all adopt the same definition. Then we would all know how to measure a fadizzle with respect to a particular observer and might be able to use this knowledge to guide decision making about the offered gambles.

3.1.5 Summary

In summary, whatever representation of uncertainty is chosen for use in engineering design, it must have both a well defined mathematical representation and a well defined operational definition of its currency, or metric (e.g. foot, meter, probability). In this dissertation, a quantity will be considered to have a sufficient operational definition if there is an unambiguous way to determine the value of that quantity in a given scenario. Also, the operational definition used to express the value of a quantity should be conveyed with the value. If no operational definition exists, communication of meaning is impossible. In this case, the representation has no use in decision making, and by extension, no use in engineering design (until an operational definition is described).

3.2 Survey of uncertainty representations

The most familiar and commonly used theory of uncertainty is the probability theory that satisfies the axioms of Kolmogorov (1956). Given its widespread use, probability will be taken as the baseline method for mathematically modeling uncertainty in engineering design. A more extensive overview and discussion of probability theory is given in Section 3.3. Initially, it is sufficient to say that probability theory has a well defined calculus and at least one well-defined, operational interpretation.

However, various other methods have been developed for representing uncertainty, including fuzzy sets (Zadeh 1965, Dai, et al. 2003), possibility theory (Zadeh

1978, Nikolaidis, et al. 2004, Mourelatos and Zhou 2005b), fuzzy measures[6] (Sugeno 1977), the transferable belief model (Smets 1990), Dempster-Shafer Evidence Theory (Shafer 1976, 1992), various theories of imprecise probabilities (Sarin 1978, Good 1983, Kyburg 1987, Walley 1991, Weichselberger 2000), and theories of (credal) sets of probabilities (Tintner 1941, Hart 1942, Levi 1974). The question addressed in this chapter is, "do these methods have significant advantages over probability theory?" This section contains a summary of the applicability of some of the most commonly considered methods, specifically fuzzy sets, possibility theory, and evidence theory. Subsequent sections contain discussions of probability theory and imprecise probability theory.

3.2.1 Fuzzy sets

Lotfi Zadeh introduced fuzzy set theory as an extension of classical set theory (1965). This section is not a detailed mathematical description of fuzzy set theory, but rather it is intended only to provide a high-level introduction to the concepts.

In classical set theory, an element's membership in a set is binary; it is either in the set (in which case the membership is 1), or it is not in the set (in which case the membership is 0). Fuzzy set theory allows for partial membership, meaning an element's membership in a set can be any real-value in the closed interval [0, 1].

---

[6] Fuzzy measures are a generalization related to monotone measure theory; the presence of the word "fuzzy" in the term "fuzzy measure" is a bit unfortunate, as it does not rely on fuzzy set theory for its rigor. Fuzzy measures are mentioned for completeness but are not necessary for understanding this dissertation and do not provide a model of uncertainty, but require refinement into particular fuzzy measures such as probability theory and possibility theory. See (Mourelatos and Zhou 2005b) for a good summary of fuzzy measure theory and the relationship to probability theory and possibility theory.

Fuzzy set theory is tightly coupled with fuzzy logic. In traditional Boolean logic, a statement is true (and has truth value 1), or it is false (and has truth value 0). Fuzzy logic extends the notion of truth and falsehood, such that a statement can "sort of" be true and take a truth value between zero and one. The connection between fuzzy logic and fuzzy sets is relatively straightforward at a high level. In classic logic, the proposition "Element b is a member of set F" would have a truth value of either zero or one, but in fuzzy logic, it can take a fuzzy truth value of any real number in the interval [0, 1]. For example, if one states that the fuzzy truth value of 0.3 is assigned to the proposition "Element b is a member of set F", then one is stating that element b is partially a member of set F, thus making the set F in some way fuzzy since the boundary of membership is not clearly demarcated.

The clearest application of fuzzy set theory is in linguistics. Language is inherent *vague*, where vague means *not clearly expressed or not having precise meaning* (Merriam-Webster 1993). Language can also be considered *ambiguous*, where ambiguous means *capable of being understood in two or more possible senses or ways* (Merriam-Webster 1993). For example, if a man is described as being tall, what is his height? What is the minimum height that he can be and qualify as tall? There is no universally accepted answer. Different people will provide different answers.

In a classroom setting, the concept of fuzzy memberships is often introduced via an experiment. Each member of the class receives a piece of paper containing a table similar to the one shown in Table 3.1. Each student is asked to place a checkmark in row 2 under each height that he or she considers "tall", and do the same in row 3 for each height he or she considers "short," such as shown in the table. The instructor then compiles the results into a second table that displays the fraction of the class that checked each box, such as shown in Table 3.2.

The values shown in Table 3.2 approximate the membership function of heights into the set "tall" and the set "short". For example, height "5 feet, 6 inches" has a membership of 0.3 in short, and a membership of 0.0 in tall. Height "5 feet, 9 inches" has a member of 0.1 in short and a member of 0.05 in tall. This is an interesting aspect of fuzzy set theory—an element can have partial membership in multiple sets, including sets that are generally thought of as mutually exclusive.

Fuzzy logic has been used with great success in the form of fuzzy controllers that generalize expert controllers (Lee 1990). A common example in this area is an inverted pendulum. In order to stabilize an inverted pendulum in the upward position, a control force must be applied to overcome disturbances and gravity. A simple expert rule might say something like, "if the pendulum is moving quickly in the counterclockwise direction, hit it hard in the clockwise direction." Another simple rule might say, "if the pendulum is moving slowly in the counterclockwise direction, hit it softly in the clockwise direction." What exactly is meant by quickly, hard, slowly, and softly? Even if these speeds and forces are defined, what about values in between? Fuzzy logic is

Table 3.1. Fuzzy set example, human height survey example response

height	4'6"	4'9"	5'	5'3"	5'6"	5'9	6'	6'3"	6'6"	6'9"	7'
tall?									√	√	√
short?	√	√	√	√							

Table 3.2. Fuzzy set example, tabulated frequencies from human height survey

height	4'6"	4'9"	5'	5'3"	5'6"	5'9	6'	6'3"	6'6"	6'9"	7'
tall?	0.0	0.0	0.0	0.0	0.0	0.05	0.15	0.3	0.7	1.0	1.0
short?	1.0	1.0	0.9	0.6	0.3	0.1	0.0	0.0	0.0	0.0	0.0

often used to interpolate between these already vaguely defined values.

Despite the value of fuzzy set theory in control applications and linguistics, it is limited in its applications to representing uncertainty in general. To date, there has been no general, operational definition given to membership values when there is no linguistic population to sample—what does it mean to say that the membership of b in set F is 0.3 rather than 0.4? How can the membership of b in set F be measured? It is true that for a particular set F, a scale could be constructed. For example, the membership function for tall was constructed using a particular method. In this linguistic context, a loosely operational definition of the membership of an element in a particular set was given; the membership of a particular height in the set tall was defined as the fraction of people in a control group who say that a man of that height is tall. No clear generalization of this definition is available, and many publications use fuzzy membership functions without making any effort to form a general interpretation.

3.2.2 Possibility theory

Possibility theory was first advanced by Zadeh (1978) as a tool for representing information expressed in terms of fuzzy measures, and has since been examined most fully by Dubois and Prade (1988). Useful overviews of possibility theory in engineering applications are given in (Nikolaidis, et al. 2004, Joslyn and Booker 2005). Three views of possibility theory have been advanced. The first is Zadeh's fuzzy set basis for possibility (1978). Another view is that possibility is the limit of the plausibility of nested bodies of evidence [for example (Klir 1992)], an aspect of Evidence Theory addressed in Section 3.2.3. A third view, given by (Giles 1982), is that a possibility is an upper probability, similar to the upper probabilities formalized by Walley (1991) and introduced in Section 3.4 of this dissertation.

The first of these interpretations (Zadeh's) is the most commonly adopted. Some authors appear to adopt the nested bodies of evidence perspective in the beginning of their analysis, only to ignore this interpretation when it becomes convenient—but incorrect—to do so. The third interpretation is not seen in the engineering literature, and it is generalized by the imprecise probability theory discussed in Section 3.4.

3.2.2.1 Basics of possibility theory

Possibility theory was first advanced by Zadeh (1978) as a tool for representing information expressed in terms of fuzzy measures. Possibility theory defines a mapping $\Pi : 2^{\Omega} \rightarrow [0,1]$ called the possibility measure, defined on a space Ω with $\Pi(A)$ for $A \subseteq \Omega$ being the degree of the possibility that A occurs (or is true, if A is a logical proposition). Possibility theory is defined such that for two sets, A and B:

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \tag{3.1}$$

$$\Pi(A \cap B) = \min(\Pi(A), \Pi(B)) \tag{3.2}$$

One of the arguments advanced in favor of using possibility theory in engineering design is the simplicity of these operations [see for example (Du and Choi 2006)]; they are concise and quick, and they involve no joint distributions or other complicated relationships. Some papers also argue that there is a clear relationship between probability and possibility theory, and that possibility theory is justified when there is little information and probability theory is useful when there is more information (for example (Du and Choi 2006)). This clearly describes a spectrum, with the two theories at opposite ends. However, this relationship is doubtful on two levels, one based on the interpretations of the models and one based on the calculi of the two models.

3.2.2.2 Lack of operational definition of possibility

First, what is a possibility? Unless possibilities are given an operational definition, it is impossible to use them consistently. Some authors have advanced a possibility-probability consistency principle (Klir and Yuan 1995, Zimmerman 1996), which states that anything that is probable must be possible. This can be interpreted to mean that the possibility of any event must be greater than or equal to the probability, or that the possibility of any event with non-zero probability is one. The latter condition is often deemed to be too conservative, and thus the former interpretation is adopted by default.

The existence of two interpretations of the possibility-probability consistency principles redirects attention of the key open question: what is possibility? What does it mean? What consequences should it have on behavior? These issues are frequently overlooked in the engineering design literature considering possibility theory, but they are essential questions. Until these questions can be answered, any reasoning in favor of possibility theory is specious.

3.2.2.3 The limitations of the calculus of possibility

The second objection focuses on the calculus of possibility. Granted, it is difficult to criticize a calculus for numbers that have no operational definition, but the argument that possibility theory is an easier way to perform calculations normally performed using probability can be countered easily. Consider two events, A and B. These two events are not independent, meaning:

$$P(A \cap B) \neq P(A) \cdot P(B) \tag{3.3}$$

Instead, the law of total probability must be used:

$$P(A \cap B) = P(B) \cdot P(A \mid B) \tag{3.4}$$

The relevant question is, how can Equation (3.4) ever reduce to Equation (3.2)? This question will be revisited from a different perspective in Section 3.2.3.8 in the context of Evidence Theory. However, the comparison of Equations (3.2) and (3.4) leads to the conclusion that possibility theory and probability theory are different, and hence any simple transform between them, such as defined in (Geer and Klir 1992) cannot be correct. The next subsection contains an expansion of this argument. However, one should note that the failure of possibility to perform probability operations is not necessarily a fundamental flaw in the theory, but it does undermine most applications of possibility theory to engineering design that have been published to date.

3.2.2.4 Lack of connection between probability and possibility

In (Geer and Klir 1992), operations for converting between probabilities and possibilities (and back) are defined. In principle, there is nothing wrong with converting one thing into another and using a different calculus for operations in the transformed domain than are used in the original domain. This is the basic idea of mathematical techniques such as Fourier or Laplace transform methods. However, these are analytical transforms with rigorous mathematical grounds. The methods proposed by Geer and Klir are merely currency transforms. By analogy, it is as if the methods tell a DM how to convert from dollars to yen and back, but with the conversion being more complicated than purely multiplicative. The transformations of Geer and Klir provide no motivation for how the operations shown in Figure 3.1 could ever, let alone always, yield the same results.

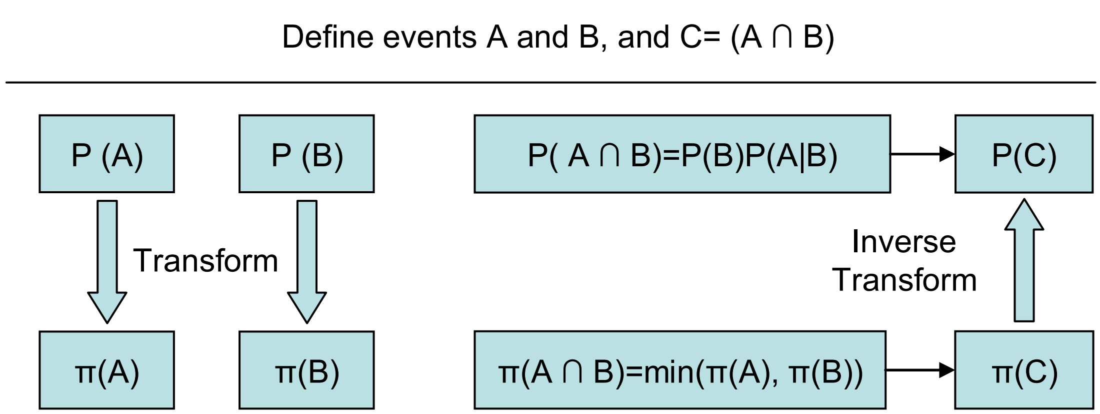


Figure 3.1: Probability-possibility transform example

The process shown in Figure 3.1 consists of defining three events, A, B, and $C = A \cap B$. The probabilities of events A and B are known, and the goal is to calculate the probability of event C. Doing this using probabilities requires the use of Equation (3.4). However, if the probabilities are converted to possibilities using methods from (Geer and Klir 1992), then it should be possible to calculate the possibility of event C using Equation (3.2) and convert this possibility back to a probability, avoiding the use of Equation (3.4). However, this result would imply that no matter what the relationship between events A and B is (no matter what P(A|B) equals), the probability of event $C = A \cap B$ remains unchanged, contradicting Equation (3.4).

3.2.2.5 Comparison between probability and possibility in a design application

Nikolaidis and co-authors (Nikolaidis, et al. 2004) have directly compared probability theory and possibility theory in the context of design against catastrophic failure. They conclude that the two methods attempt to minimize failure in "radically different ways", especially when more than one failure event is involved. They provide one particularly revealing example:

> Consider a highly redundant system that fails only if many unfavorable events occur simultaneously. The system failure region is usually small, which tends to make the probability of failure small, whereas the possibility of failure can be still high. For example, the probability of failure of a system of *n* nominally identical, independent components connected in parallel decreases exponentially with *n*, whereas its possibility is equal to the possibility of failure of a single component, regardless of *n, which is counterintuitive*. In this case, possibility theory is conservative.

The behavior described is a direct consequence of Equations (3.2) and (3.4), and that $P(A|B) = P(A)$ when events A and B are independent. .

According to arguments such as in (Geer and Klir 1992), it should be possible to map the probabilities of failure of each component into possibilities of failure for each component (an expanded version of the process in Figure 3.1). Each of these possibilities could then be transformed back into probabilities, but this is mostly irrelevant. The real question is whether such transformations hold through mathematical operations in each domain, which it is relatively clear from the Nikolaidis and co-authors example (Nikolaidis, et al. 2004) is in general not true. Until such difficulties are resolved, any work drawing relationships between probabilities and possibilities is highly questionable. Nikolaidis et al. emphasize this point in one direction, but as described above, the converse clearly holds, too:

> *Possibility and probability calculi are fundamentally different*. We cannot simulate the results of possibility calculus using probability calculus by properly selecting the parameters of the probabilistic models.

The inability of possibility theory to reach an accord with probability theory is not a reason to reject the theory in its entirely. It could be that possibility theory is capable of representing a particular type of uncertainty that cannot be modeled accurately with probability theory. However, this returns to the question of what is possibility? The existing literature has not reached a clear operational definition of possibility. As such, uses of possibility theory in design as the only model of uncertainty are certainly unjustifiable.

3.2.2.6 Positive results of using possibility theory in design, existing literature

Nikolaidis and co-authors (Nikolaidis, et al. 2004) and others (Mourelatos and Zhou 2005b, Du and Choi 2006) claim to have achieved reasonable results by using possibility theory in design. However, of the papers known to the author, only Nikolaidis and coauthors have shown this in a rigorous manner with clearly defined assumptions. The remainder of the known work, such as (Du and Choi 2006), are fundamentally flawed, primarily due to the irresponsible or undefined transforms between probability and possibility measures.

The results of Nikolaidis and coauthors (Nikolaidis, et al. 2004) indicate that possibility theory can sometimes lead to much more conservative or less conservative results than probability theory, depending on the circumstances:

> *Possibility can be less conservative than probability in risk assessment of systems with many failure modes*. Possibility-based methods tend to underestimate the risk of failure of such systems, especially if the number of modes is large. *Possibility tends to yield more conservative estimates of the risk of failure for systems for which many unfavorable events have to occur simultaneously in order to produce failure*. An example is a parallel system.

Nikolaidis and coauthors reach the final conclusion:

> If we have enough information about uncertainties and accurate predictive models, then probability is advantageous. On the other hand, *when making design decisions under limited information or using crude predictive models it may be useful to consider both the probability and possibility of failure of a system*.

Phrased differently, the conclusion is basically that when a large amount of information is available, probability theory is the best approach. On the other hand, when information is very scarce, both a probability theory based analysis and a possibility theory based analysis should be performed, and the design based on the most conservative result.

If a “more conservative of the two” design approach is adopted, the resulting design may be overly conservative when the possibility theory model yields the more conservative result. This is a consequence of possibility not having a clearly defined operational definition. If it is not clearly defined, there is no way to determine if the result with possibility of failure of x is too conservative, not conservative enough, or just right.

3.2.2.7 Summary of the value of possibility theory as an uncertainty model

Until a clear, operational definition is given that is applicable in engineering design applications, there is little value in forming design methods based on possibility theory. In the extreme, with no operational definition given to possibilities, an engineer can assume whatever values he or she wants, as the values have no intrinsic connection to the environment or his or her own beliefs and behaviors.

3.2.3 Evidence theory

Evidence theory, also referred to Dempster-Shafer theory, was introduced by Glenn Shafer (1976) when he expanded the work of his advisor, Arthur Dempster (1967). However, its roots go back to George Hooper, James Bernoulli, and Johann-Heinrich Lambert (Shafer 1978, 1986).

3.2.3.1 Basic set operations

Evidence theory takes the n possible outcomes (or states of nature) and forms a mutually exclusive and exhaustive set $\{a_1,...,a_n\}$ of n-outcomes. This set is called the frame of discernment Θ, and the members of the set are called focal elements. This is not any different from the probability formulation of n mutually exclusive and exhaustive events $\{E_1,...,E_n\}$ forming the sample space S. The difference is the way in which the evidence or probability is assigned across these outcomes. Rather than assigning probabilities or belief purely to individual mutually exclusive events, evidence theory assigns belief to any element in the power set of outcomes.

For example, consider the case with $n = 3$. Then $\Theta = \{a_1,\ a_2,\ a_3\}$, the full list of subsets in the power set is $\{a_1\}$, $\{\ a_2\}$, $\{a_3\}$, $\{a_1,\ a_2\}$, $\{a_1,\ a_3\}$, $\{a_2,\ a_3\}$, $\{a_1,\ a_2,\ a_3\}$. Depending on the available evidence, each of these subsets will be supported to some degree. For example, there could be evidence that supports both $\{a_1\}$ and $\{\ a_2\}$ but not $\{a_3\}$ and also does not distinguish between $\{a_1\}$ and $\{\ a_2\}$. Thus the evidence is for the subset $\{a_1,\ a_2\}$ and is assigned using the basic belief mass function.

3.2.3.2 The basic belief mass function

The basic belief mass function, also called an m-value and sometimes a basic probability assignment, is the amount of evidence specifically and exclusively for a particular subset of the frame of discernment. Similar to probabilities, the sum of the m-values across all subsets must equal one. Thus for all subsets A of the frame Θ:

$$\sum_{A \subseteq \Theta} m(A) = 1$$

The value $m(A)$ for a given set A can be thought of as the fraction of all relevant and available evidence that supports a particular member of the frame that belongs to set A but to no particular and distinguishable subset of A. If no evidence is available, then $m(\{a_1,\ a_2,\ a_3\}) = 1$; all that is known is that the outcome is in the frame of discernment.

The additivity rule of evidence theory is shown in Equation (3.5) for $n = 3$.

$$m(a_1) + m(\ a_2) + m(a_3) + m(a_1,\ a_2) + m(a_1, a_3) + m(a_2,\ a_3) + m(a_1,\ a_2,\ a_3) = 1 \quad (3.5)$$

For comparison, the additivity rule for probability theory is shown in Equation (3.6).

$$P(a_1) + P(a_2) + P(a_3) = 1 \quad (3.6)$$

Unlike basic belief mass, probability must be assigned to a particular focal element or basic event. Basic belief can be assigned to sets of outcomes $\{a_1,\ a_2\}$, $\{a_1,\ a_3\}$, $\{a_2,\ a_3\}$, $\{a_1,\ a_2,\ a_3\}$ as well as the individual outcomes $\{a_1\}$, $\{\ a_2\}$, and $\{a_3\}$ to which probability theory is limited. This added expressiveness can more accurately reflect ambiguity.

3.2.3.3 The belief function

The belief function is different from the basic belief mass function. The belief function essentially states (for a particular outcome) the minimum possible amount of belief in the outcome that could remain after all ambiguity is fully resolved. Mathematically, the belief in subset A is given as:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{3.7}$$

For example, the $Bel(\{a_1\}) = m(a_1)$, since the only subsets of $\{a_1\}$ are $\{a_1\}$ and $\{\varnothing\}$, and $m(\varnothing) = 0$. However, $Bel(\{a_1, a_2\}) = m(a_1) + m(a_2) + m(\{a_1, a_2\})$.

3.2.3.4 The plausibility function

In contrast to the belief function, the plausibility function is the maximum amount of belief that could support an outcome or subset if all ambiguity is resolved in a way that supports that outcome or subset. Mathematically, if $\neg A$ is defined as the set complement to the set A, then

$$Pl(A) = \sum_{A \cap B \neq \varnothing} m(B) = 1 - Bel(\neg A) \tag{3.8}$$

For example, $Pl(\{a_1\}) = m(a_1) + m(\{a_1, a_2\} + m(\{a_1, a_3\} + m(\{a_1, a_2, a_3\})$ because the evidence that supports any set containing a_1, such as $\{a_1, a_2\}$, may in fact actually support $\{a_1\}$, but the analyst is unsure. The second equality in Equation (3.8) reflects that the plausibility of A is equal to the fraction of the evidence that is not explicitly against A. Evidence that is not explicitly for $\neg A$ could yet be allocated to A, whereas evidence specifically against A can never support A. In other words, a plausibility function measures the degree to which evidence does not contradict the belief in an outcome.

3.2.3.5 The ambiguity function

As explained above, the belief in A, $Bel(A)$),is a minimum or lower bound on the evidence for A, and the plausibility, $Pl(A)$, is an upper bound. The lower bound of $Bel(A)$ and upper bound of $Pl(A)$ reflect the extremes of how the ambiguity can be resolved—either entirely for A or against A respectively. The ambiguity function is defined as the difference between the plausibility of A and the belief in A.

$$\mathit{Ambiguity}\text{ in }A = Pl(A) - Bel(A) \tag{3.9}$$

3.2.3.6 Combination of evidence

One of the original contributions of Dempster and Shafer was a system for combining evidence from multiple sources. The original rule proposed by Dempster (1967) is a generalization of Bayes' rule. This rule illustrates some of the fundamental challenges in combining evidence or subjective information. Dempster's Rule, as it is called, states that two independent basic-belief functions (m_1 and m_2) are combined according to Equation (3.10) (Shafer 1986, Sentz and Ferson 2002).

$$\begin{aligned}
&m_{12}(A) = \frac{\sum_{B\cap C=D} m_1(B) * m_2(C)}{1-K} \text{ when } A \neq \varnothing \\
&\text{such that } m_{12}(\varnothing) = 0 \text{ and} \\
&\text{where } K = \sum_{B\cap C=\varnothing} m_1(B) * m_2(C)
\end{aligned} \tag{3.10}$$

Dempster's Rule of Combination has come under significant scrutiny because of its seemingly irrational results in the following example, originated by Lotfi Zadeh (1984b) and restated by Sentz and Ferson (2002):

> Suppose that a patient is seen by two physicians regarding the patient's neurological symptoms. The first doctor believes that the patient has either meningitis with a probability of 0.99 or a brain tumor, with a probability of 0.01. The second physician believes the patient actually suffers from a concussion with a probability of 0.99 but admits the possibility of a brain tumor with a probability of 0.01. Using the values to calculate the $m(\text{brain tumor})$ with Dempster's rule, we find that $m(\text{brain tumor}) = Bel(\text{brain tumor}) = 1$. Clearly, this rule of combination yields a result that implies complete support for a diagnosis that both physicians considered to be very unlikely.

For the purposes of this dissertation and for comparing uncertainty models, it is important to note that the combination of evidence is a separate issue from the representation of uncertainty and arises in all uncertainty models (Kacprzyk and Fedrizzi 1990, Yager, et al. 1994a, Oberkampf and Helton 2002). It is therefore important not to throw the baby out with this bath water; just because Dempster's rule of combination is sometimes invalid does not mean that Evidence Theory is unsuitable as an uncertainty model.

3.2.3.7 Interpretations of basic belief function

Some people think of evidence theory as merely upper and lower bounds on probabilities. While the results of applying evidence theory are a lower measure and an upper measure, they are in general not bounds on probabilities. Shafer renamed the lower measure "belief" in order to avoid this confusion (1992), because Dempster had originally introduced a system of "lower probabilities" (Dempster 1967). Shafer makes it clear he does not believe that belief and plausibility should be viewed as bounds on probability, because they are not always constructed using probability measures. While there are cases in which the bounds are on probabilities, the power of evidence theory is in its generality. Evidence theory can represent many types of evidence, including complete experimental frequency data (such as probabilities), sparse experimental results, or expert opinions. Evidence theory does not define what belief is, but merely provides the most general framework for set operations on any units of measurement. While the units of belief do not need to be probabilities, they can be probabilities, as discussed in the next subsection.

3.2.3.8 Probability theory and possibility theory as special cases of Evidence Theory

Probability and possibility theory can both be viewed as special cases of evidence theory. The key is the limitation of the sets and subsets allowed in the problem. Consider a case with three focal elements. In general evidence theory, these three elements $\{a_1\}$, $\{a_2\}$, and $\{a_3\}$ are assumed to be disjoint, meaning $\{a_i\} \cap \{a_j\} = \varnothing,\ i \neq j$. Probability makes a similar assumption. In Evidence Theory, basic belief mass can be assigned to the power set of these focal elements, meaning all of the sets $\{a_1, a_2\}$, $\{a_1, a_3\}$, $\{a_2, a_3\}$, in addition to the individual singletons $\{a_1\}$, $\{a_2\}$, and $\{a_3\}$. Probability theory only allows probability to be assigned to the individual outcomes $\{a_1\}$, $\{a_2\}$, and $\{a_3\}$.

For convenience, the notation is now defined such that for a set b, $|b|$ represents the number of disjoint focal elements in the set. For example, $|\{a_1\}| = 1$ and $|\{a_1, a_3\}| = 2$. If it is the case in evidence theory that $m(b) = 0$ for all sets b such that $|b| > 1$, then evidence theory reduces to probability theory, assuming basic belief is given the same operational definition as probability.

Possibility theory can also be viewed as a special case of evidence theory. Possibility theory is equivalent to a set of nested focal elements. A set of focal elements is nested if (using the example of three focal elements), it is true that the elements can be ordered such that $c_1 \subseteq c_2 \subseteq c_3$. Taking the power set of evidence theory, if it is defined that $c_1 = \{a_1\}$, $c_2 = \{a_1, a_2\}$, and $c_3 = \{a_1, a_2, a_3\}$, then this relationship holds. If the set $C = \{c_1, c_2, c_3\}$ is defined, then evidence theory reduces to possibility theory if $m(b) = 0$ for all b such that $b \not\subset C$ and if possibility is defined as plausibility.

For example, $\Pi(c_1 \cup c_2) = \max(\Pi(c_1), \Pi(c_2))$. In Evidence Theory, the plausibility of $c_2 = \{a_1, a_2\}$ is given by Equation (3.11), which is found recalling Equation (3.8) and the restriction that $m(a_2) = 0$, since $\{a_2\}$ is not a subset of C.

$$Pl(c_2) = \sum_{c_2 \cap b \neq \varnothing} m(b) = m(a_1) + m(a_1, a_2) \tag{3.11}$$

Since $c_1 \subseteq c_2$, then $\Pi(c_1 \cup c_2) = \Pi(c_2)$. With possibility defined as plausibility, then $\Pi(c_1) = Pl(c_1)$ and $\Pi(c_2) = Pl(c_2)$ and $\Pi(c_1 \cup c_2) = \max(Pl(c_1), Pl(c_2))$. Since $Pl(c_1) = \sum_{c_1 \cap b \neq \varnothing} m(b) = m(a_1)$, from Equation (3.11) it is clear that $Pl(c_1) \leq Pl(c_2)$, and hence $\max(Pl(c_1), Pl(c_2)) = Pl(c_2)$.

In summary, restricting the focal elements to be nested sets and defining the units of possibility to be units of plausibility reduces Evidence Theory to possibility theory. Similarly, restricting focal elements to disjoint sets and defining probability as basic belief reduces the general Evidence Theory to the specific probability theory. In this context, the stark differences between probability theory and possibility theory are revealed in yet another way, once again calling into question transformations between the two (Geer and Klir 1992) and existing applications of possibility in engineering design such as (Du and Choi 2006).

3.2.4 Summary of uncertainty representations

This section has discussed three alternatives to probability theory that have been advanced in the literature, and in some cases used successfully in specialized applications. However, they are all flawed in a basic way—they have no clearly established operational interpretation for the units of the measures. Until such interpretations are firmly established and demonstrated, there is little use in considering the methods for the complex decision making under uncertainty required in engineering design.

3.3 Probability theory

Probability is the most generally used model of uncertainty (Winkler 1996), and some authors would argue is the only acceptable way to represent and propagate uncertainty, as cited in (Cox 1946, Lindley 1982b, Cheeseman 1988, Ferson and Ginzburg 1996, Lindley 2000).

The term "probability" is used informally in common speech to mean many things. It also has several meanings within the uncertainty research community. However, when most people refer to probability in a mathematical sense, they are referring to one formalization based on Kolmogorov's axioms (1956). This is the probability theory introduced in most undergraduate probability and statistics courses, and this probability theory is clearly an acceptable formulation of the mathematics of aleatory uncertainty (inherent randomness). Other valuable references in foundations of probability are DeFinetti (1974, 1975, 1980) and Savage (1972).

3.3.1 Axioms of probability

Probability theory is formalized as follows. In an experiment, the set of all possible outcomes is the sample space S. The outcomes are called events E_i, such that $E_i \subseteq S$. A probability assigns a number $P(E_i)$ to the event, such that $P(E_i)$ is a measure of the "probability" that event E will occur. The assignments of the probability function $P(\cdot)$ must satisfy three axioms

1. Non-negativity: for any event E_i,

$$P(E_i) \geq 0 \tag{3.1}$$

2. Normalization:

$$P(S) = 1 \tag{3.2}$$

3. Additivity: If E_1, E_2,….are disjoint events (disjoint subsets of S), then:

   Finite additivity: for a finite collection of n events:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P\left(E_i\right) \tag{3.3}$$

   Countable additivity: for an infinite collection of events:

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P\left(E_i\right) \tag{3.4}$$

The first two axioms are primarily matters of convenience that have become standard convention. A consequence of the axioms is that for disjoint events,

$$\sum_{i=1}^{n} P\left(E_i\right) = 1 . \tag{3.5}$$

Furthermore, rearranging

$$P(E_j) = 1 - \sum_{\substack{i=1 \\ i \neq j}}^{n} P\left(E_i\right) . \tag{3.6}$$

Equivalently, this last equation can be written as $P(E_i) = 1 - P(\neg E_i)$. Stated in words, this means that the probability of event E_i occurring is one minus the probability of event E_i not occurring. This is because $\{E_i\} \cup \{\neg E_i\} = \{S\}$ and due to the normalization axiom of Equation (3.2). For clarity, it may help to recall Equation (3.2) and note that for disjoint events:

$$\{E_i\} \cap \{E_j\} = 0 \text{ for all } i \neq j . \tag{3.7}$$

3.3.2 Basic calculus of probability theory

The basic calculus of probability for the set operations of union and intersection are given in Equations (3.8) and (3.9). Since these relationships are quite familiar to most engineers and are covered in basic probability and statistics courses and books, such as (Devore 1995, Hines, et al. 2003), no further details are given here.

$$\mathrm{P}(\mathrm{A} \cap \mathrm{B}) = \mathrm{P}(\mathrm{B}) \cdot \mathrm{P}(\mathrm{A}|\mathrm{B}) \tag{3.8}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{3.9}$$

3.3.3 Interpretations of probability

Most engineers are familiar with the mathematics of probability, but many have not been formally exposed to the competing interpretations of probability. The philosophical arguments for or against different interpretations can be quite passionate. The most important interpretations are the classical, logical, frequentist, propensity, and subjective interpretations. Each of these is described in the following.

3.3.3.1 Classical interpretation

The classical interpretation of probability was the original interpretation and formalization of probability. Probability has its roots in games of chance and was first developed when gamblers asked Pierre Fermat and other mathematicians to calculate exact probabilities in games of chance (Carnap 1950). Classical probability was most formally developed in the writings of Jacob Bernoulli (1713), Thomas Bayes (1764), and Pierre Simon de Laplace, with Laplace's treatise (1951) being the most complete. A fundamental assumption in classical probability is what is often called Laplace's Principle of Insufficient Reason or Principle of Indifference. This principle states that all outcomes are equally probable (Ferson and Ginzburg 1996, Winkler 1996).

The classical view is more than just an interpretation of probability, it is more fundamentally a definition. Laplace (1951)defines probability as follows:

> The theory of chance consists of reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible.

Various objections to this interpretation have been raised, not the least of which is the impracticality of requiring equi-probable events. Various paradoxes and inconsistencies have also been raised, including Bertrand's paradoxes. As explained by von Fraassen (1989), the following example illustrates a paradox. A factory produces cubes with sides of lengths between 0 feet and 1 foot. The question asks, "What is the probability of a given cube having a side-length between 0 and 0.5 feet?" The immediate answer appears to be ½. However, the scenario can be recast equivalently as follows. A factory produces cubes with surface area per face between 0 and 1 square feet. The question now becomes, "What is the probability that a given cube has face-area of 0.25?" The immediate answer now appears to be 0.25. Is this reasonable?

As stated, the two scenarios are actually identical. A cube with side-length between 0 and 1 foot has face-area between 0 and 1 square feet. A cube with side-length between 0 and 0.5 feet has face-area between 0 and 0.25 square feet. So depending on how the question is framed, the classical interpretation of probability can yield different probabilities for exactly the same problem. Should the probabilities of the same events vary depending on how they are named or described? This inconsistency is a definite limitation of the interpretation and a clear problem for guiding decision making.

A version of the classical interpretation is often adopted in practice when there is little or no information available about the probabilities. With no information suggesting an uneven allocation of probabilities, each event is allocated an equal probability. For example, consider the roll of a six-sided die. Nothing is known about the fairness of the die, so how should the probabilities be allocated? According to the Principle of Insufficient Reason, the probabilities should be allocated equally, as in Equation (3.10).

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6 \tag{3.10}$$

Notice that this is the same result as in the case of full knowledge that the die is fair. In variations of classical probability, it is impossible to distinguish complete ignorance from complete knowledge of truly equally probable events. These two extremes are modeling the same way, meaning the true state of knowledge is not expressed. In cases where nothing is known, the knowledge is severely overstated, and the model is not conservative.

Carnap argues that the Principle of Insufficient Reason should at most be used in cases of obvious symmetry in the problem, such as the faces of a die or the sections of a roulette wheel (Carnap 1950). Otherwise, the probability of unknown events is rarely even, such as an example Carnap cites from Jeffreys (1939) in which he argues that according to the Principle of Insufficient Reason, the probability that there is life on Mars is 0.5, because there is (or at least was not in 1939) neither sufficient evidence for nor against the proposition.

An argument frequently made for the Principle of Insufficient Reason is that it is the most "uncertain" of all distributions, since it has the highest Shannon entropy (Shannon 1948). However, in this example it is clear that this is a very special case of the world—a case in which the game is fair. It is arguable that this is not the null-hypothesis of most people. It seems more reasonable that people assume a game of chance is unfair unless there is evidence that it is fair, since many people who run games of chance are out to make a profit by luring in un-savvy gamblers.

3.3.3.2 Frequentist interpretations

The frequentist interpretation, also known as objective or relativist, is based on the notion of relative frequencies of outcomes. von Mises (1939) and Reichenbach (1949) were among the founders of the frequentist interpretation. Under a frequentist interpretation, a probability represents the limit of ratio of times that one outcome occurs compared to the total number of outcomes in an endless series of identical trials (Winkler 1996).

For example, an experiment is repeated for n identical and independent replications. The number of times, $N(E_i)$, the experiment's outcome is a particular outcome E_i is counted. The relative frequency is defined as $N(E_i)/n$. In practice, this ratio tends to stabilize as n gets large, such that the relative ratio approaches a limiting value. The frequentist interpretation of probability takes this limiting ratio as the probability of E_i, denoted $P(E_i)$ as in Equation (3.11).

$$P(E_i) = \lim_{n\to\infty}(N(E_i)/n) \tag{3.11}$$

The frequentist interpretation is considered to be an objective interpretation because it is based on experimental outcomes rather than interpretations of individual observers. For experiments such as die tossing, the outcomes are objective. The result is indisputably 1, 2, 3, 4, 5, or 6; there is no subjective interpretation of the result. In theory, two observers will record the same result for a particular replication and thus calculate the same relative frequencies.

The frequentist interpretation fails when there are no examples from which to determine frequencies (Pate-Cornell 1996). For example, consider the design of a satellite. A current design decision depends somewhat on the type of fasteners that will be chosen during detail design. What is the ratio of times that a particular bolt will be used to fasten two components of a satellite? This particular satellite design task is novel—it has never been done before. How can designers construct relative frequencies about something that has never happened before? They might be able to look at all past satellites and construct frequencies that way, but are those trials really identical since each satellite has a novel purpose?

The frequentist interpretation defines probability as the limiting relative frequency of an outcome in a sequence of identical and independent experiment replications. In general, engineers turn to probability theory in the very cases in which the events or experiments of interest are not repeatable. The design process itself is not repeatable. Even if designers follow a rigorous, systematic design process with stable objectives, the results will not be identical due to human factors. For example, the same idea will not necessarily come along at the same time. Other examples of non-repeatable events to which people often try to assign probabilities are election results, sports events, and contract negotiations (Devore 1995, Hajek 2003). In the real world, none of these exact events happens more than once.

Philosophically, even if an event can be repeated, a particular set of replications is still only a sample of the true frequency. No matter how large the number of replications n is, the relative frequency can only be calculated to the granularity of $1/n$ (Hajek 2003). Going even further, for a particular n, the set of outcomes can be thought of as one experiment, in which the experiment itself is to take n samples. This one replication of a calculation of the relative frequency is clearly not a good estimate of the true relative frequency. Although in theory the relative sample frequency approaches the true relative frequency as the sample size goes to infinity, an infinite sample size is impossible to acquire in practice due to resource constraints. Consequently, engineers will always face imprecision in their characterizations of the frequentist probabilities.

In most engineering situations, agreement on convergence is not an issue for large data sets. The problem is more often that only limited data exists. If only 10 data points are available, assessment of a limiting frequency involves imprecision. For a fair coin-tossing experiment, the odds of getting 7 heads out of 10 flips are 0.117. This means that when the true probability is 0.5, there is a 0.117 chance of estimating it to be 0.7. Given 10 data points, engineers cannot be sure whether they know the true probability, but they do have some information that can guide decision making. The open question is how to model this information and the related uncertainty appropriately in order to support decision making.

As the number of replication decreases, the knowledge of the probabilities decreases and becomes more uncertain. In the extreme case already noted in the context of novel satellite design, there may be no existing data about a particular experiment. From a purely frequentist perspective, in this scenario nothing can be said about the probabilities. As suggested earlier, engineers may look at similar designs and allocate probabilities based on those outcomes, but those trials are not identical. The engineers are taking a leap of faith when they extrapolate from past designs to current designs. Why should the new design have the same probabilities? If the new design has different probabilities, how close might they be to past designs? These questions cannot be answered objectively or in a frequentist framework.

The basic lesson is that the true frequency of an can never be known, and the assignment of a frequentist probability requires certain assumptions to be made, such as assuming or judging whether frequency has converged "close enough." Such assumptions are ultimately subjective in nature (Gelman, et al. 1995). Other researchers recognize this subjectivity and suggest a purely subjective interpretation of probability.

3.3.3.3 Subjective interpretation

This section is begun with a few historical notes. The subjective interpretation of probability has its foundations in work by Borel (1909) and Ramsey (1926). Lindley (1982b) provides an excellent overview of the subjectivist view, both in terms of probability and decision making. The greatest extensions of the theory were made by Savage (1972)[the most general] and Anscombe and Aumann (1963) [the most elegant and complete, though less general]. The subjective interpretation is sometimes referred to as the Bayesian interpretation, though the interpretation is entirely separate from Bayes rule and is applicable even when not using a Bayesian analysis approach such as in (Berger 1985).

A strict subjective interpretation of probability asserts that there are no true or inherent probabilities (or that they cannot be known), but rather probabilities are constructed from individual opinions. In this sense, probability is a degree of belief. Each person has their own belief in a given situation. These beliefs are based on the individual's knowledge, assumptions, preferences, and biases. A person's beliefs are reflected by his or her betting behaviors. Based on the work of de Finetti (1980) and as discussed in (Hajek 2003), belief has the following paraphrased meaning:

> Your degree of belief in outcome E is p if and only if p units of utility is the price, known as a *fair price*, at which you are indifferent between buying or selling a bet that pays 1 unit of utility if outcome E occurs, and 0 units of utility if E does not occur. Assuming a linear (risk neutral) utility function and no endowment effects in this problem, p then represents the individual's probability that outcome E will occur.

A more complete and theoretically correct definition would phrase things in terms of different lotteries of payoffs. For example, consider the following example adapted from Cooke (2004).

- Two events are defined:

  S: Slovakia wins the next Olympic Men's Hockey Gold Medal.

  U: The USA wins the next Olympic Men's Hockey Gold Medal.

- Two lottery tickets are available:

  LS: worth $1000 if S occurs, worth $10 otherwise.

  LU: worth $1000 if U occurs, worth $10 otherwise

A person (Jane) is offered one of these lotteries, and she may choose whichever she prefers. The truth proposition "Jane's degree of belief in S is at least as great as her degree of belief in U" is operationalized (or observable) as the event "Jane chooses lottery LS." In Savage's formalization (1972), Jane's choice of lottery LS over lottery LU also implies that her subjective probability that Slovakia wins the gold, $P(S)$, is no smaller than her subjective probability that the USA wins the gold, $P(U)$. Mathematically, $P(S) \geq P(U)$.

Notice that a subjective interpretation of probability is also a behavioral interpretation. In a behavioral interpretation, a probability model must have implications concerning an individuals behavior (Walley 1991). This is a reasonable requirement for probability models used in decision making; if a probability has no connection to an individual's behavior in decision making, how can a probability guide decision making?

There are several complications with the subjective interpretation of probability. First, a person's willingness to enter a bet depends on more than just probabilities. The utility a person derives from an outcome depends on his or her preferences and his or her current state of wealth. Thus choices reflect not just beliefs but also preferences. This is the reason that Savage axiomatized preferences and belief (utility and probability) simultaneously (Savage 1972). Second, people are inherently bad at assessing probabilities (Tversky and Kahneman 1974, Kahneman, et al. 1982). Finally, there are other objections such as people having multiple prices at which they would enter a bet, and that your fair price p also measures your belief in the bet actually being paid, given that you win (Hajek 2003).

Despite some limitations, the subjective interpretation is often used in practice. For example, in the satellite bolt design problem of the previous section, the subjective interpretation of probability is probably better than the frequentist interpretation. Even though frequency data for the novel design is not available, an experienced designer might be able to make a subjective judgment about historical frequencies of past designs. The result would be subjective probabilities regarding the novel design.

For example, assume past historical data shows a log-normal distribution with mean 5 and variance 1 for a particular dimension. The designer also knows that the new, high-strength, light-weight composites used in this design will reduce the stress seen on the bolt below past averages. He may combine this knowledge with the past frequencies and come to a conclusion on how he would be willing to bet on the outcome of the novel design. He may conclude that he would bet according to probabilities defined by a lognormal distribution with mean 4 and variance 1.

One of the primary arguments against a frequentist interpretation and for a subjective interpretation of probability is the absence of truly repeatable events, especially in practical problems. For example, the probability that Team A beats Team B in a basketball game has no real meaning under a frequentist interpretation, because that event—that particular game—will occur exactly once. In this context, the notion of a long term frequency, and even random events, is meaningless (de Finetti 1974).

However, people are often still willing to bet on the outcome of a game. This betting behavior reveals information about people's beliefs, which in turn can be connected to probabilities using subjective probability theory. In this way, subjective probabilities could in theory be found for any problem, though in practice this requires the expenditure of resources and may be impractical.

3.3.3.4 Other interpretations of probability

The logical interpretation, also known as the necessary view, of probability was first proposed by Keynes (1921) and later advocated by Jeffreys (1939) and Carnap (1950). Carnap has argued for logical probabilities in addition to statistical (e.g. frequentist) probabilities, while Keynes and Jeffreys, who were writing when the frequentist interpretation was in its own infancy, claimed that logical probabilities were the only legitimate probabilities.

Under the logical interpretation, the probability of a hypothesis is uniquely determined given a particular body of evidence (Walley 1991). This interpretation has some inherently attractive qualities to it—such as implying a universally correct answer. However, the logical interpretation has met with many challenges, including how anyone could ever know the correct or "logical" probability for a particular body of evidence. Consequently, it is generally regarded as dead (Cooke 2004).

A final interpretation that deserves mention is the propensity interpretation. Like the frequentist and logical interpretations, the propensity interpretation asserts that probabilities are properties of the world rather than of the decision maker. Thus, they are entirely objective in nature. According to Popper (1957, Popper 1959b), a probability is an inherent tendency—a propensity—of a system to produce a particular result with some frequency. The propensity is a fundamental attribute of the system, rather than a reflection of other empirically observable attributes. The propensity itself is unobservable, except in a frequentist type setting. As such, there is no way to empirically test the value of a propensity, so it is impossible to develop an operational theory of propensities and probabilities for use in decision making.

3.3.3.5 Adopted interpretation of probability

Given the failures of the classical, frequentist, logical, and propensity interpretations of probability to be useful in engineering design decisions, a subjective interpretation is the most applicable of existing theories. In comparison to a frequentist interpretation, a subjective interpretation is applicable to a broader class of problems, as it is not limited to repeatable events. The subjectivist interpretation advocated here is not as strict as the traditional views (Lindley 1982b), because it admits imprecisely known subjective probabilities, as described in Section 3.4. The traditional school claims that by definition, subjective probabilities are known to a decision-maker, because they are his or her beliefs. Section 3.4 describes the advantages of an interpretation that acknowledges the practical difficulties (Weber 1987, Walley 1991, Groen and Mosleh 2005) in arriving at a precise characterization of such beliefs. Fundamentally, the interpretation advocated is subjective in that probabilities are defined in terms of beliefs and betting behavior.

The use of subjective probabilities is often troubling to many engineers and scientists, as they often have a preference for the objective. Intuitively, there should be some connection between available evidence and beliefs, but this connection is not identical for all people, and there is rarely a way to enforce consistency. Walley (1991) suggests a rationalist interpretation that requires probabilities to be consistent in certain ways with the evidence, without requiring that they be uniquely determined (as was required by the logical interpretation). For example, subjective probabilities should be consistent with observed relative frequencies when such frequencies are available. However, the exact correlations between these observed frequencies and the subjective probabilities are not well defined. Consequently, the rationalist interpretation is incomplete and hence not suitable for adoption as a strict approach, but more as a vaguely defined heuristic.

Fundamentally, there is nothing wrong with using subjective probabilities. After all, preferences are clearly subjective and vary from person to person. When beliefs (probabilities) and preferences are combined in decision making (as is the case with expected utility theory), why should the probabilities be objective when the utilities, and hence the optimal decision, is subjective anyway? For the reasons addressed through out this Section (3.3.3), the subjective interpretation of probability is adopted throughout this thesis. In certain problems, such as the example problem presented in Chapter 5, these subjective probabilities appear very "frequentist" in nature, because it is assumed that the decision-maker's beliefs are based on observed frequencies and statistical analysis, as explained in more detail in that chapter.

3.3.4 Traditional statistical decision theory and utility theory

The most commonly adopted method for making decisions with uncertainty represented by probabilities is traditional statistical decision theory. In traditional statistical decision theory (Pratt, et al. 1995), utility analysis, as formalized by von Neumann and Morgenstern (1944), is used for making decisions under uncertainty. von Neumann and Morgenstern originally derived their theory using objective probabilities, but Savage (1972) later extended utility theory to subjective probabilities, a theory referred to as subjective expected utility (SEU). For extended discussion of expected utility theory, see references (Raiffa 1968, Fishburn 1982) for the foundations or references (Thurston 1990a, Thurston 2001, Scott 2004) for summaries in the context of engineering design.

In general, utility expresses preference—more preferred decision outcomes are assigned higher utility values. Utility theory has been studied extensively by economists and decision theorists, and there continues to be an increasing interest in and discussion of applying utility methods to engineering problems (Thurston 1990a, Hazelrigg 1998, Fernández, et al. 2001, Thurston 2001, Scott 2004). If chosen correctly, utilities reflect the decision-maker's preferences, even under uncertainty. By applying the expected value operator, the decision-maker weights all possible outcomes according to their likelihood of occurring, and then chooses the action that maximizes the expected utility.

Utility theory is a normative theory. Essentially, it described how a rational individual should act. The notion of rationality is encoded into the axioms of the theory. As a normative theory, it does not need to be confirmed by all observations or descriptions of human behavior. There is actually substantial evidence that humans do not act rationally according to utility theory (Tversky and Kahneman 1974, Kahneman, et al. 1982). The goal of a normative theory is not to replicate current practice, but to provide an ideal for which to strive.

A primary assumption of utility theory is that probabilities, alternatives, and preferences are fully known by the decision maker. In practice, a decision maker could lack knowledge about all of these (Weber 1987). In this dissertation, the emphasis is on probabilities, which leads to the question as to whether probability theory can appropriate capture the lack of knowledge, or imprecision (defined in Chapter 2), faced in engineering design decisions.

3.3.5 Ability of probability theory to represent imprecision

A traditional probability distribution is precise in nature. There is exactly one probability associated with every event. For example, one can write $P(A) = 0.6$, or $P(A) = 0.7$, but not both. There is no way to express a lack of knowledge or imprecision in the true value. Writing $P(A) = (0.6 \text{ or } 0.7)$ would lead to a violation of the normalization and additivity axioms, Equations (3.2) and (3.3). For example, if $P(A) = (0.6 \text{ or } 0.7)$, then $P(\neg A) = (0.3 \text{ or } 0.4)$. This could lead[7] to the expression $P(A) + P(\neg A) = 0.7 + 0.4 = 1.1$, a violation of the axioms.

Consider the distribution shown with a solid line in Figure 3.2. This distribution is a best-fit normal distribution based on 15 samples of a random process. Since this distribution is based on a small number of samples, there is a great deal of imprecision in addition to the irreducible uncertainty from the random process. How can this imprecision be represented using probability?

In order to use traditional probability theory, a designer is forced to eliminate imprecision, ignore imprecision, or confound imprecision with a different aspect of uncertainty. Eliminating imprecision requires the designer to expend resources to acquire more information, thus increasing the costs of the design process. Ignoring imprecision involves overstating the true current state of information by making assumptions. This is equivalent to forcing a decision-maker to make an exact statement or choice, even if he or she has not yet reached an exact belief of judgment (Weber 1987).

---

[7] Assuming that just the probabilities are considered without carrying additional information, such as dependencies and the other relationships, through the calculations.

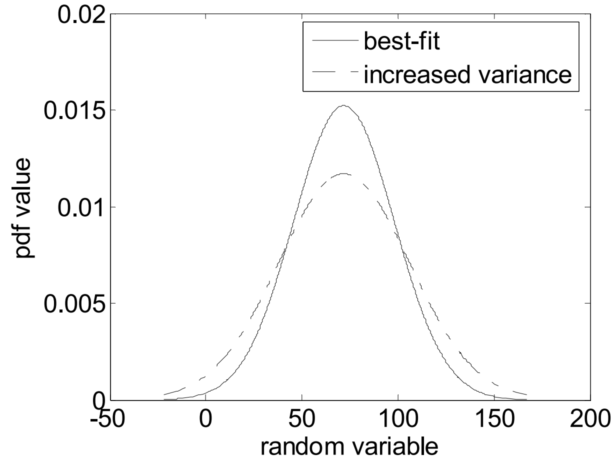


Figure 3.2. Example distribution adjusted for imprecision

Common practice is to force imprecision into the representation by proxy by increasing the variance of the distribution, such as shown with the dashed line in Figure 3.2. In a way, this change does increase the total uncertainty (if uncertainty is equated with Shannon entropy (Shannon 1948)), but it confounds two different aspects—imprecision and irreducible uncertainty—of the problem, and therefore makes it difficult to draw useful insights from the resultant distribution (Helton 1994). For example, there is no way to distinguish the dotted distribution, which is now an imprecise representation of the true (but unknown) distribution, from a distribution that is precisely known to be the dotted distribution. An additional example further illustrates this point.

When traditional probability theory is used to represent both irreducible uncertainty and imprecision, no information for problem formulation and information collection decisions is provided, as illustrated with the following example. Consider a scenario in which two experimenters perform the same experiment and find the best-fit normal distributions shown in Figure 3.3. Which is the better fit? Actually, one of the experimenters performed 20 tests, and thus one of the distributions represents data from a sample size of 20. The other represents a sample size of only 5. Which is which? Traditional probability theory gives no indication as to how much information a distribution is based on. By extension, it gives no indication as to how useful additional information could be.

Bayesian analysis (Berger 1985) takes a slightly different approach. It starts with a prior distribution and update this distribution as information is received. The priors are chosen to be non-informative, meaning they contain as little information as possible. As information is acquired, the distributions become more informative. As was the case with the example in Figure 3.2, variance is again used as a proxy for imprecision in this method.

The choice of priors in Bayesian analysis is not unique (Walley 1991). Empirical Bayes methods have the goal of using Bayesian analysis without fully specifying the prior distribution or its parameters (Robbins 1955, Efron and Morris 1973). Bayesian sensitivity analysis (Berger 1985) considers a set of priors. Each prior (or input) distribution yields a different posterior (or output) distribution. The decision maker is thus left with a set of distributions to consider. If the decision is robust (e.g. does not change) across this set of priors, then the uncertainty in the choice of an appropriate prior does not matter for that particular decision. However, the problem remains regarding any reflection of the amount of information on which the distribution is based, as was

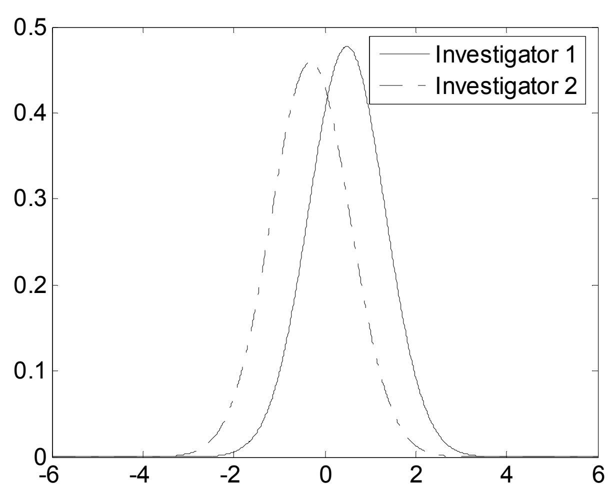


Figure 3.3. Example distribution adjusted for imprecision

illustrated using Figure 3.3.

In conclusion, if probability can capture both imprecision and irreducible uncertainty, it does so in a way that confounds the two, making independent insights into either impossible. This in turn makes it difficult to make efficient decisions regarding information collection, as there is no indication as to the information already in the distribution. In the extremes, there is no way to distinguish a case with no knowledge, e.g. a uniform prior distribution, from a case of precise knowledge, e.g. a perfectly known uniform distribution. Consequently, in order to consider imprecision explicitly and separately from irreducible uncertainty, it is necessary to consider a different model of uncertainty. The remaining core of this dissertation (the remainder of this chapter plus most of Chapter 5 through Chapter 7) presents such models—called imprecise probabilities and the sub-theory of probability bounds analysis—and then establishes the practical value of using such models that in general separate imprecision and irreducible uncertainty.

3.4 The theory of imprecise probabilities

Some alternatives to precise probabilities, such as fuzzy sets, possibility theory, and the some interpretations of Dempster-Shafer Evidence Theory, abandon the notion of probability entirely. This opens them up to significant criticism, at least in part due to their unfamiliarity to practicing engineers, but more importantly due to their lack of a clear operational or behavioral interpretation that was discussed earlier in this chapter.

On the other hand, theories based on imprecise probabilities (Sarin 1978, Good 1983, Kyburg 1987, Walley 1991, Weichselberger 2000) or sets of probabilities (Tintner 1941, Hart 1942, Levi 1974) are an extension of traditional probability theory and therefore have clear operational and behavioral interpretations. For this reason, attention is now focused on imprecise probabilities as a general uncertainty model. In the remainder of this section, imprecise probabilities are motivated, defined, and advanced as a useful model of uncertainty.

3.4.1 Motivation for imprecise probabilities

The general motivation for imprecise probabilities is that the more evidence on which a probability estimate is based, the more confidence a decision-maker can have in it. Thus, the imprecision in the probabilities should be expressed explicitly in order to signal the appropriate level of confidence to ascribe to them. For example, a decision maker would like to place more confidence in the distribution in Figure 3.3 that is based on 20 samples than the one based on just 5 samples. However, the distributions do not reflect such information.

It is more straightforward to motivate and introduce imprecise probabilities using discrete events rather than continuous random variables. Adapting an example from Walley (1991), consider the exercise of determining the probability that a tossed thumbtack lands pin-up. Three experimenters perform this exercise, as follows:

- Experimenter A is in a hurry and does not even look at the thumbtack. Experimenter A employs a non-informative prior distribution (e.g. uses the principle of indifference or insufficient reason) and assumes that the probability of the tack landing pin-up is equal to the probability it lands pin-down, thus ascribing a probability of 0.5 to both outcomes.

- Experimenter B tosses the thumbtack 10 times and gets 6 pin-ups. Experimenter B's estimated precise probability of the tack landing pin-up is thus 0.6.

- Experimenter C tosses the thumbtack 1000 times and gets 400 pin-ups. Experimenter C's estimated precise probability of the tack landing pin-up is thus 0.4.

If the three experimenters are three analysts that could provide the DM with information, which analyst would the DM prefer to hire? Because Experimenter C's estimate was based on more data, it is more precise than Experimenter B's estimate. Experimenter A's estimate was based on no data, so it does not seem reasonable to place much confidence in it. Nevertheless, the precise probability estimates of 0.5, 0.6, and 0.4 appear equally credible. By not expressing the imprecision in these estimates, one is arbitrarily eliminating it by assuming precision that has no justification in the available evidence. This problem can be overcome by allowing analysts to state imprecise probabilities.

3.4.2 Definitions of upper and lower previsions and probabilities

A full discussion of Walley's formalization of imprecise probabilities (1991) is outside the scope of this dissertation, but a summary is necessary. In the context of the probability interpretations discussed in Section 3.3.3, Walley's imprecise probabilities are subjective. Walley defines upper and lower probabilities as special cases of upper and lower previsions (de Finetti 1980). In simple terms, a DM's lower prevision is the highest price at which the DM is sure he or she would buy a gamble, and the upper prevision is the lowest price at which the DM is sure he or she would buy the opposite of the gamble (which is equivalent to selling the original gamble). If the upper and lower previsions are equal, then they jointly represent the DM's fair price for the gamble, the price at which the DM is willing to take either side of the gamble. The existence of a fair price leads to precise probabilities, as described in Section 3.3.3.3. Previsions are equivalent to probabilities if the stakes of the gamble are one unit of currency, such as one dollar or one util.

Adopting Walley's notation, lower previsions are denoted $\underline{P}$, and upper previsions $\overline{P}$. The set of possible states of the world is denoted Ω. A gamble X represents a bounded, real-valued function on Ω that is interpreted as an uncertain reward. Every gamble, such as shown in Figure 3.4, has a buyer and a seller, and the consequences of the gamble affect each different. The seller sells the gamble to the buyer for some *price*, with the price being transferred from the buyer to the seller at the time the gamble is bought. A gamble X will pay the buyer a reward of x if a particular event happens, and a reward of zero otherwise. A gamble X will cause the seller to incur a penalty of x if the event occurs and a penalty of zero otherwise.

A DM's lower prevision $\underline{P}(X)$ for gamble X represents the maximum price at which the DM is willing to buy the gamble X. In other words, the DM will readily purchase the gamble for any price less than $\underline{P}(X)$. More specifically, if $\underline{P}(X) = \mu$, then μ is the supremum buying price for which it is asserted than the gamble $X - \mu$ is still desirable to the decision maker. The quantity $X - \mu$ is considered because if the payout of a gamble is the uncertain quantity X and the DM pays μ to enter the bet as a buyer, then the net payoff to the DM is $X - \mu$, where $X = x$ if the event occurs, and $X = 0$ otherwise. A rational, risk-neutral DM will only enter into this gamble if the expected payoff of $E[X - \mu]$ is greater than zero. For eliciting previsions, it is noted that small gambles (gamble for which x is much smaller than the DM's wealth) are more suitable.

The connection to probabilities is made as follows. First, the possible payoff of the gamble X is defined as $X = x = 1$ if a particular event, say A, occurs, and $X = 0$ otherwise. Then the lower prevision on gamble X becomes the lower probability of event A. For such a bet, the expected payout to the buyer of $E[X - \mu]$ is equal to

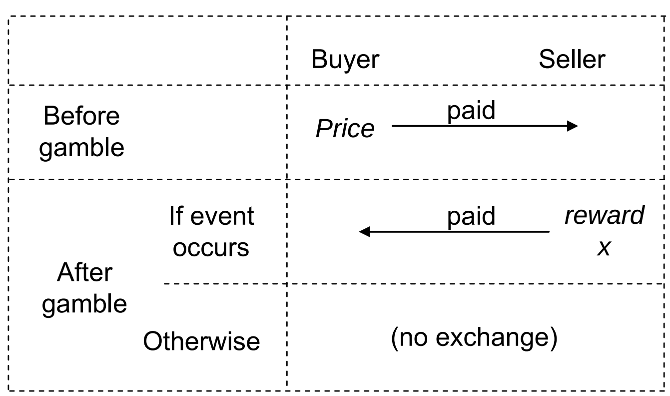


Figure 3.4. Anatomy of a gamble: money transfers

$E[X] - \mu = p(A) - \mu$, since μ is a constant and $E[X] = p(A) \cdot 1 + p(\neg A) \cdot 0$. As such, a rational, risk-neutral DM will require that $p(A) \geq \mu$ in order to enter the bet, and thus the highest price μ that the DM is willing to pay for the gamble represents the highest lower-bound on the DM's probability that event A occurs.

For a particular lower prevision $\underline{P}$, Walley defines the conjugate upper prevision $\overline{P}$ as $\overline{P}(X) = -\underline{P}(-X)$. $\overline{P}(X)$ is the infimum selling price for the gamble X, meaning that if $\overline{P}(X) = \alpha$, then α is the smallest price for which the DM is willing to sell the gamble X. Note that when the DM sells the bet for α, the net payoff to the DM is $\alpha - X$, since the DM now gains α when the bet is sold, and pays out—a loss—the uncertainty quantity X when the gamble is realized. The definition of an upper probability is made by extension from an upper prevision in the same way that a lower probability was an extension of a lower prevision (i.e., in terms of a gamble with payoff of $X = x = 1$ to the buyer if a particular event, say A, occurs, and $X = 0$ otherwise).

If a DM spends the time and effort to collect complete evidence and fully elicit his or her beliefs, then the imprecise probabilities will collapse into precise probabilities, such that $\alpha = \mu$ and $\underline{P}(X) = \overline{P}(X)$. Traditional probability requires that $\underline{P}(X) = \overline{P}(X)$ (and hence $\alpha = \mu$), meaning the DM is indifferent between selling and buying the gamble at price $\alpha = \mu$, the so-called fair price. If the gamble is offered at a price higher than $\alpha = \mu$, the DM will sell the gamble. If the gamble is offered at a price lower than $\alpha = \mu$, the DM will buy the gamble. If the bet is offered at exactly $\alpha = \mu$, then the DM is willing to take either side.

There is a difference between being willing to take either side of a gamble and neither side. The case of taking neither side is not allowed in traditional probability theory, which requires that a DM be able to state, given a price, his or her willingness to the buy or sell the gamble; the DM is not allowed to refrain from gambling entirely. In Walley's theory of imprecise probabilities, at any price ρ such that $\mu < \rho < \alpha$, the DM will neither sell nor buy the gamble, but rather will refrain from gambling altogether.

3.4.3 Axioms of coherence and avoidance of sure loss

Walley begins his formulation of imprecise probabilities with the same fundamental notion of rationality as de Finetti (1974, 1980)—avoidance of a sure loss (also known as a Dutch Book). Walley's axioms of coherence assure that if a DM's imprecise probabilities satisfy them and the DM makes decisions consistent with these probabilities, then the DM is not subject to a sure loss. His primary deviation from axioms of precise probabilities such as de Finetti's is the allowance for a range of indeterminacy—prices at which the DM will not enter a gamble as either a buyer or a seller. By not entering a gamble at prices between his or her upper and lower previsions, the DM will not subject himself or herself to a sure loss.

Walley's axioms for coherent lower previsions are as follows:

P1: $\underline{P}(X) \geq \inf X$ (accepting sure gains)

P2: $\underline{P}(\lambda X) = \lambda \underline{P}(X),\ \ \lambda > 0$ (positive homogeneity)

P3: $\underline{P}(X+Y) \geq \underline{P}(X) + \underline{P}(Y)$ (super-linearity)

The axioms for coherent upper previsions are defined as follows:

P1a: $\overline{P}(X) \leq \sup X$ (accepting sure gains)

P2a: $\overline{P}(\lambda X) = \lambda \overline{P}(X),\ \ \lambda > 0$ (positive homogeneity)

P3a: $\overline{P}(X+Y) \leq \overline{P}(X) + \overline{P}(Y)$ (sub-linearity)

As noted earlier, it is also required that $\underline{P}(X) = -\overline{P}(-X)$, or equivalently $\overline{P}(X) = -\underline{P}(-X)$. Combined, the axioms also require that $\inf X \leq \underline{P}(X) \leq \overline{P}(X) \leq \sup X$. It is also true that $\underline{P}(A) + \overline{P}(\neg A) = 1$, meaning the total probability of events in the universe must add to one. Walley asserts and then demonstrates that any set of upper and lower previsions satisfying these axioms will avoid sure loss.

The avoidance of sure loss is a weaker condition than coherence, but it is a necessary condition. As Walley writes, "Incoherence [i.e. a violation of his axioms] reflects a kind of ignorance about the consequences of Your [e.g. the DM's] judgments, whereas failure to avoid sure loss reflects a fundamental irrationality or error in judgment." For reference, Walley continues (in his Section 2.5-2.6) to develop consequences of avoiding sure loss and basic properties of coherent previsions. He then covers the generalization to probabilities (in his Section 2.7) and the relationship to traditional probabilities (in his Section 2.8).

3.4.4 Eliciting and assessing

Basic definitions, properties, and axioms of imprecise probabilities are a starting point, but they are of no use to designers unless there are ways to determine the probabilities. In Chapter 4 of his book, Walley tackles the issue of assessing and eliciting probabilities. According to Walley, elicitation is the process by which beliefs (pre-existing behavioral dispositions) are measured, through explicit judgments and choices. Assessment is the process by which probabilities are constructed from the available evidence.

Walley presents some basic elicitation procedures. These begin with direct judgments of desirability (comparative judgments) and proceed to more complicated judgments. Walley also addresses the issue of natural extension, or the process of updating the models that results from your previous elicitations with evidence from new elicitations, which can be viewed as part of the assessment process. For example, an increase in a DM's lower probability for an event has particular consequences on his or her upper probability for the complement of that event, since $\underline{P}(A)+\overline{P}(\neg A)=1$.

Despite the rigor of the foundations, the elicitation and assessments of imprecise probabilities may be a large obstacle to the widespread use of imprecise probabilities. This challenge is similar to that of utility theory (von Neumann and Morgenstern 1944), in which rational behavior under uncertainty is defined under the assumption that utility functions and precise probabilities can be found. While procedures for eliciting utilities are defined (von Neumann and Morgenstern 1944, Keeney and Raiffa 1993), the use of these procedures in practice can be challenging and resource intensive (Seepersad 2001).

3.4.5 Imprecise probability distributions

In the previous sections, imprecise probabilities were described in terms of discrete probabilities. However, imprecise probability theory can be extended to probability distributions. A probability density function $f(x)$ is defined such that $\int_a^b p(x)dx = P(a \le x \le b)$, or the probability that x is between a and b. The event that x is between a and b will now be defined as event A, and thus $\int_a^b p(x)dx = P(A)$. Then upper and lower bounds on the density function could be defined as $\int_a^b \underline{p}(x)dx = \underline{P}(A)$ and $\int_a^b \overline{p}(x)dx = \overline{P}(A)$. However, this is somewhat limiting in that it established relationships using bounding distributions.

A more general approach is to think of a set of distributions. If one assumes that the information available for a particular decision at a given point in time is given by I, then one can define the set $P(I)$ as the set of all probability density functions that are *consistent* with the given information I. The notion of consistency is somewhat loose. Something is consistent with the available information if the available information does not rule it out. For example, if the evidence is only that $x < 5$ and x is limited to positive integers, then $P(x = 4) = 1$ and $P(x = 4) = 0.3$ are both consistent with the available information, but $P(x = 6) = 0.2$ is inconsistent with the available information.

All of the distributions in the set (all $p_i(x) \in P(I)$) are distributions that the DM cannot rule out as being the true distribution, just as in the discrete case the DM cannot rule out any probabilities in the interval $[\underline{P}(x), \overline{P}(x)]$. In the discrete case, the DM will only buy or sell gambles for which the expected payoff is positive for all probabilities in this interval. In the continuous case, the DM will only buy or sell gambles for which the expected payoff is positive for all distributions in the set. If the DM accepted a gamble that had a negative expected payoff for some distribution (say $p_j(x)$) in the set, it could turn out that that distribution ($p_j(x)$) is the true distribution and the DM would have accepted gamble with an expected loss, which contradicts rational, risk-neutral behavior.

3.4.6 Discussion of objections to imprecise probabilities

As noted earlier, at a price between the upper and lower previsions, the DM is not willing to enter the bet on either side, at least not without collecting more information and updating his or her previsions. One objection to the use of imprecise probabilities is that they can lead to such indeterminacy of action during decision-making. That is, given imprecise probabilities, there may not be a single, clear "best" solution according to standard decision theories.

This argument is countered by stating that if the available evidence does not clearly suggest a particular course of action, then the representation of this evidence should not arbitrarily pretend that it does. An approach that demands precise probabilities necessitates an arbitrary resolution of the indeterminacy. Such an approach does not differentiate well-grounded probabilities (such as Experimenter C's in the tack-tossing example of Section 3.4.1) from arbitrary ones (such as Experimenter A's). By admitting imprecise probabilities, one can abstain from these arbitrary judgments during analysis and support better decision-making, as follows.

It is true that in order for a single solution to be chosen, any indeterminacy will need to be resolved. However, the explicit representation of imprecision in the characterization of uncertainty allows for the direct management of the imprecision in the context of a decision. For example, if a decision is not sensitive to the current level of imprecision, a robust decision can be made using arbitrary rules[8] such as arbitrary choice (Walley 1991), Γ-maximin (Berger 1985), or the Hurwicz criterion (Arrow and Hurwicz 1972). On the other hand, if the decision is sensitive to the existing imprecision, one can decide to collect more information (such as more tack tosses in the earlier example), perhaps managing the set of alternatives according to policies of E-admissibility (Levi 1974) or maximality (Walley 1991). These topics are discussed at greater length in Chapter 8.

Another frequently levied objection to imprecise probabilities is the Dutch Book Argument that they are irrational (de Finetti 1974, Lindley 1982a, Walley 1991). The general idea of a Dutch Book is that if a DM's probabilities violate certain rules, a group of bets can be constructed, all of which the DM is willing to accept, but the combination of which results in a sure loss; the DM will lose money under any outcome. This argument is often presented in favor of precise probabilities and the axioms of Kolmorogov (1956). However, as described in Section 3.4.3, Walley (1991) presents axioms of coherence for imprecise probabilities that also avoid sure loss. Key in this proof is the assumption that a decision is not mandated; that is, indecision is an acceptable conclusion. The idea is that if the available information does not support a clear choice, any choice made is arbitrary.

Arbitrary decision policies may encounter problems because it is difficult to employ these arbitrary rules consistently across multiple decisions. The use of arbitrary

[8] These rules are arbitrary in nature, meaning the distinction is not made based on rational guidance; according to rational principles and the available information, the decision-maker's preference is indeterminate.

choice across a sequence of decisions may in fact lead to the possibility of a Dutch Book and sure loss. However, the use of imprecise probabilities in situations in which choice is not mandated and arbitrary choice is avoided cannot lead to a Dutch Book. Policies for decision making in the presence of imprecise information are discussed in Chapter 8.

3.4.7 Computational limitations of imprecise probabilities

In general, computing with imprecise probabilities is computationally demanding, especially in the case of imprecise probability distributions. Most methods are based on mathematical programming. However, one can think of an abstraction of what is required. In this abstraction, the most basic, brute-force means for computing involves a second-order, or double-loop Monte Carlo process. Recall that a set of distributions $P(I)$ that are consistent with the available information is defined (see Sections 3.4.5 and 3.6.2 for more). A second-order Monte Carlo procedure contains two loops, one nested inside the other. In the outer loop, the set of distributions $P(I)$ is sampled, where each sample is a particular distribution $p_i(x) \in P(I)$. In the inner loop, a more traditional Monte Carlo sampling is performed by sampling points from the currently sampled $p_i(x)$. These points are used to estimate the integral of interest for the analysis (e.g. expected value of a function) for the chosen $p_i(x)$. The process then repeats to the outer loop, in which a different distribution is sampled, such as $p_j(x) \in P(I)$. By repeating this process, a set of output values corresponding to the sampled input distributions can be found; each output corresponds to one of the consistent input distributions in $P(I)$.

It is important to note that the outer loop is not being used to estimate a robust statistical parameter (such as mean or median). If that were the case, then the set $P(I)$ could be sampled heavily and the parameter estimated by averaging the results. However, the goal is to determine the set of output values corresponding to the possible distribution $P(I)$. It is thus in theory necessary to consider all of them, aggregating them into a summary such as an interval that shows the range of values that can occur for the output given the imprecision in the inputs.

This double loop process is very computationally expensive. For example, assuming there are N distributions in the set $P(I)$, the outer loop must be repeated N times in order to be exhaustive. The inner loop requires M samples from each distribution, where M is often a number between 100 and 50,000 (Ferson and Ginzburg 1996). Defining F as the cost of one evaluation of the function or model of interest, the cost of the double loop process is $O(N \cdot M \cdot F)$.

In general, there could be an infinite number of distributions in $P(I)$. In this situation, the double loop process is prohibitively expensive. There is no way to systematically narrow a general set $P(I)$ to a finite number of representative distributions to consider. However, one could place restrictions on the set $P(I)$, essentially creating a sub-set of imprecise probability theory. This is essentially what is done in probability bounds analysis (PBA) (Ferson and Donald 1998), a method of modeling uncertainty that is described in detail in Chapter 4.

PBA puts relatively mild restrictions on $P(I)$ but leads to computations of the order $O(K^2F)$, where K is the number of discretization bins used in the computational algorithm[9] and K is generally less than 1000 and often less than 100 (Ferson and Ginzburg 1996). Probability bounds analysis essentially limits $P(I)$ to the set of probability density functions $f_i(x) \in P(I)$ for which the corresponding cumulative distribution functions $F_i(x) = \int_{-\infty}^{x} f_i(t)dt$ are such that $\underline{F}(x) \le F_i(x) \le \overline{F}(x)$—in other

[9] There are actually multiple algorithms and heuristics for computing in PBA. These results use the original method, called *dependency bounds convolution*. Additional information on computing in PBA is provided in the next chapter.

words, distributions for which the cumulative distribution function (CDF) is bounded from below by a particular CDF defined as $\underline{F}(x)$ and from above by the CDF defined as $\overline{F}(x)$. A much more complete discussion of PBA is presented in Chapter 4, but this brief discussion has motivated why PBA is a reasonable model; it is much more convenient for computing and only slightly less expressive than general imprecise probability theory.

3.5 Hierarchical uncertainty models

One can view imprecise probability theory as a type of hierarchical uncertainty model. In an hierarchical uncertainty model [about which a good review is presented in (de Cooman and Walley 2002)], uncertainty is expressed about another uncertainty. The most basic hierarchical uncertainty model is a *second-order uncertainty* model. These methods apply when it is assumed that there is some ideal uncertainty model describing a phenomenon of interest, but there is uncertainty as to which model is the correct model. For example, a set $P(I)$ of distributions that are consistent with the available information was defined in Section 3.4.5. Each of these distributions is a particular model of the first-order uncertainty. One of the distributions in $P(I)$ is the true or ideal distribution, but the DM is unsure which. In a second-order uncertainty approach, the DM would model this uncertainty about the uncertainty, perhaps even using as second-order probability, as described in the following.

3.5.1 Second-order probabilities

In a second-order probability model in the Bayesian community, probability measures are defined over a set of probabilities or other probability measures (Good 1980, Berger 1985). One simple model is just to assume that each of the probability distributions in $P(I)$ is equally likely to be the true distribution. In this example, the second-order uncertainty is modeled using a uniform distribution; the DM has assigned particular probabilities to each of the probability distributions, which is can be called a second-order probability model (Utkin and Augustin 2003).

In theory, if second-order probabilities are embraced, then third-order, fourth-order, and so on probabilities can also be embraced. Each higher-order probability describes the likelihood of some particular lower-order probability being the correct description of that order probability. Each level becomes less precise, because if a higher-order probability were more precise, it could readily be collapsed to a lower order. For example, if the second-order probabilities were zero for all $p_j(x) \in P(I)$ except for some $p_i(x)$ for which the second-order probability were one, then the first order probability is known precisely to be $p_i(x)$ and there is no need for higher order probabilities.

In the context of decision making with expected value maximization (e.g. expected utility maximization), the use of second-order probabilities provides little benefit. The result of taking the expectation over the two probabilities distributions is still a single point estimate of the value. Consequently, the second-order probability model could be collapsed into some first-order probability model that is equivalent (Walley 1991, Utkin 2003).

From an elicitation perspective (meaning the process of determining subjective probabilities), it may be useful to consider second-order probabilities even though they could collapse into a single first-order probability. For example, it may provide a useful way of thinking about the problem. It may be easier for a DM to first define a set of candidate probabilities, and then express relative beliefs about them. These "second-order" beliefs really reflect a "first-order" belief about the problem, since the higher order distributions could be collapsed to a first-order probability, but the process may be a useful construct for eliciting this information. These comments are speculative in nature, and as such the value of higher-order probability models in elicitation is an area for future work. However, de Cooman and Walley (2002) note that the related Bayesian hierarchical model has been found useful, especially when one individual (the "modeler") is modeling the beliefs of another individual (the "subject").

3.5.2 Imprecise probabilities as a second-order uncertainty model

An interval is a specific model of uncertainty. An interval provides upper and lower bounds on some value, but provides no information about where in the interval a value falls or is likely to fall. An interval clearly the best model of at least one source of uncertainty; specifically, the uncertainty related to a non-detection (Ferson, et al. 2002a).

More specifically, Ferson and coauthors describe how many measuring devices have a particular detection threshold. In an environmental risk assessment, a DM may be interested in the concentration x of some chemical in a sample. However, more measurement devices have some threshold—a detection limit below which the device will not detect the presence of the chemical.

If the detection limit is D and the device does not detect the chemical in a sample, then all that can be said about the sample is that the concentration x is in the interval $[0, D]$. Nothing can be said about the probability of where x falls in the interval. A rigorous analysis must be based on the entire interval, such that all of the points are considered, because all of those points could be the true one. In order to be sure to consider the true one, all must be considered.

Imprecise probability theory adopts a similar view towards the interval of probabilities that are consistent with the available information. Because the DM cannot be sure where in this interval the true probability lies, his or her actions must be consistent with all probabilities in the interval.

An interval is a particular type of set, specifically a bounded neighborhood of real numbers. Imprecise probability theory works in this more general setting, in which the set of probabilities $P(I)$ that are consistent with the available information is considered.

For example, if upper and lower probabilities are stated for some uncertain quantity X, then $[\underline{P}(X), \overline{P}(X)]$ represent the set of probabilities that are consistent with the available information.

The key distinction between imprecise probability theory and second-order probability models is that no probabilities are assigned over the members of the set $P(I)$. This is considered the most basic expression of imprecision; the DM lacks the information necessary to identify exactly which probability is the most appropriate or the true probability. However, other researchers have explored the issue of expressing non-interval second-order uncertainties.

3.5.3 Other hierarchical uncertainty models

Researchers have proposed other hierarchical uncertainty models, including fuzzy probabilities (Watson, et al. 1979, Zadeh 1984a, Pan and Yuan 1997), second-order possibility distributions (Walley 1997), the "confidence-weighted" upper and lower probabilities of Nau (1992), and the "reliability measures" of Gärdenfors and Sahlin (1982, 1983). More recently, de Cooman and Walley (2002) have developed a model that includes aspects of these other models, but also has a behavioral interpretation.

de Cooman and Walley note that both a Bayesian hierarchical model (second-order probability) and a possibilistic hierarchical model can be useful, concluding that [quoting a reference to (Walley and De Cooman 2001) that appears in (de Cooman and Walley 2002)]:

> In summary, the Bayesian hierarchical model appears to be appropriate when the modeler has extensive information about a Bayesian subject's uncertainty, whereas the possibilistic hierarchical model seems to be appropriate when the modeler has little information about the subject's uncertainty or when her information is of a special kind

which commonly arises from vague probability assessments.

de Cooman and Walley (2002) note that the normal situation falls between those two extremes, and as such a more general model is needed. For example, Walley (1991) Section 5.10 contains a general hierarchical model in which coherent upper and lower previsions are used to represent both first-order and second-order uncertainties. Upper and lower previsions are the core measures in the theory of imprecise probabilities, and as such, it appears that the theory summarized in Section 3.4 is a rigorous and general starting point for building a generalized uncertainty model that captures both imprecision and irreducible uncertainty.

3.6 A general mathematical representation of decisions under uncertainty

Given the conclusion that imprecision is an important component of uncertainty and that imprecise probabilities are a promising model of uncertainty, it is useful to formulate a mathematical model of the design problem. While this model is not widely used in the dissertation outside of Chapter 6, the consideration of its construction provides valuable insight into the nature of the problem.

The model is built in three stages, beginning with a traditional, precise model. In the development of the model, certain concepts are more easily grasped in the continuous case, while others are more easily understood in terms of the discrete, finite case. The models developed in this subsection are not meant to be complete or universally applicable. They are presented primarily to form a context for the discussion in the remainder of this thesis.

3.6.1 Precise model

A DM has a set of alternative actions $A = \{a_1, ..., a_m\}$ (assumed finite for simplicity) from which to choose a single action. The chosen action will be confronted by some uncertain future state of the world. This state is one of the (assumed finite) set $S = \{s_1, \ldots, s_r\}$ of all possible states of the world. Each state of the world has a probability $p(s_j)$ associated with it. For a given pair of an action and a state, the DM will receive a single consequence $g(a_i, s_j)$. The DM's preferences are modeled by a utility function $u(g)$. According to von Neumann-Morgenstern utility theory (von Neumann and Morgenstern 1944), the DM should choose the action a^* that maximizes the expected utility, given by Equation (3.12) for the discrete case.

$$E[u] = \sum_{j=1}^{r} p(s_j) \cdot u(g(a_i, s_j)) \tag{3.12}$$

If the state of the world is expressed with a continuous variable, Equation (3.13) can be used instead.

$$E[u] = \int p(s) \cdot u(g(a_i, s)) ds \tag{3.13}$$

The normative model of utility theory is appropriate if the probabilities, utility functions, and consequences are known perfectly. However, the preceding chapters have shown that such assumptions are not always reasonable.

3.6.2 Basic imprecise model

The model presented in this section is based on the general model for decision making with incomplete information developed by Weber (1987). The first change from the basic precise model of Section 3.6.1 is the allowance for imprecise probabilities. In the continuous case, this means that rather than being represented by a single distribution $p(s)$, the probability density of the state of the world is given by a class, or credal set (Levi 1980), of distributions. In the discrete case, this would mean that $p(s_i)$ would equal a set (e.g., interval) of values rather than a single value.

More generally, if one assumes that the information available (which includes information about possible actions, the states of the world, and the DM's own beliefs and preferences) for a particular decision at a given point in time is given by I , then one can write that $p(s) \in P(I)$, meaning that $P(I)$ is the set of all probability density functions that are consistent with the given information I, as was done in Section 3.4.5. Furthermore, a DM can have incomplete information about more than just probabilities. While this dissertation focuses on probabilities, the decision model derived here will be more general.

In addition to imprecise probability information, a DM may not have full knowledge of his or her preferences, and consequently may not be able to identify an appropriate utility function. It is therefore useful to define $U(I)$ as the set of all utility functions consistent with the available information. Note that a particular utility function will naturally be a function of the design alternatives, but the admissibility of a particular function into $U(I)$ is theoretically independent of the available alternatives. The admissibility of a function into $U(I)$ is purely a function of the DM's preferences and the available information about them. It may be possible in practice to further limit the set $U(I)$ based on information about the alternatives. For example, if no alternative produces dangerous air contaminants, then there is no reason to model the DM's preferences for air pollution into the utility model.

One can define $G(I)$ as the set of all consequences that are consistent with the available information. The utility functions are then defined such as $u(g(a,s))$, where $u \in U(I)$ and $g(a,s) \in G(I)$. A consequence $g(a,s)$ is a function of the action taken, a, and the state of the world, s. Possible causes for the existence of only imprecise knowledge about the consequences of a particular action-state pair include the use of simplified models or an incomplete understanding of physical phenomena, as was described in Section 2.4.

3.6.3 Generalized imprecise model

The preceding model is reasonably complete assuming that the state of the world and all of the consequences of an action can each be summarized with a single parameter. In many cases, engineers do not develop their analysis models in these terms, but rather consider multiple types of consequences and multiple uncertainties.

3.6.3.1 Multiple types of consequences

More generally, engineers are concerned with multiple consequences. These consequences are often measured in quantities such as mass, cost, and performance. It is therefore necessary to generalize the consequence function into a vector of consequences, $\vec{g}(a,s)=\{g_1(a,s),...,g_h(a,s)\}$. In this model, $G(I)$ becomes the set of consequence vectors that are consistent with the available information. In the general model, it is important to note that the individual $g_i(a,s)$ do not necessarily represent analytical functions. In fact, the form of the $g_i(a,s)$ may change depending on the alternative chosen. For example, the function $g_i(a,s)$ that calculates the volume of a container depends on the geometry, so if alternative a_1 is a sphere and alternative a_2 is a cube, clearly $g_i(a_1,s_1)$ will be evaluated using a different formula than $g_i(a_2,s_1)$ will be. It is thus crucial to think of the g_i 's as consequence types rather than functions.

3.6.3.2 Multiple utilities

The expansion to multiple consequence types requires a change in the utility formulation. Specifically, the problem is now a multi-attribute utility problem (Keeney and Raiffa 1993, Thurston 2001, Scott 2004). In this model, it is assumed that a utility function can be defined for each consequence, such that if g_i represents the mass of the design, a utility function $u_i(g_i)$ can be created that reflects the DM's preference for mass. It is assumed that the individual utilities can be aggregated into a single utility measure, for example by using the additive independence form shown in Equation (3.14). For a given action a and state s, where the k_i's are weighting parameters.

Multiplicative relationships are also possible. See (Keeney and Raiffa 1993, Thurston 2001, Scott 2004) for more information.

$$u(g(a,s)) = \sum_{i=1}^{h} k_i \cdot u_i(g_i(a,s)) \tag{3.14}$$

3.6.3.3 *Multiple uncertain parameters*

The preceding models assume that the uncertain state of the world is captured by a single parameter s. Engineers rarely think in these strict terms but rather view things in terms in of multiple uncertain parameters. For example, the mass of a rectangular block is given by Equation (3.15).

$$mass = density \cdot length \cdot width \cdot height \tag{3.15}$$

Due to manufacturing imperfections and measurement errors, none of the parameters can be determined precisely. In a strict sense, any combination of uncertainties in these four parameters represents a single state of the world, but engineers generally think of them as separate (though not necessarily independent) events. Engineers may specify intervals, probabilities distributions, or even p-boxes (see Chapter 4) for each parameter. In order for the mathematical model to describe such approaches, the notion of a state must be extended.

A state is now defined to consist of a vector of uncertain quantities, such that $s = \{{}_1x, \ldots, {}_nx\}$. For example, perhaps $s = \{mass, volume, strength, ..\}$ The uncommon pre-subscripts are used to avoid confusion with common notation in which x_i represents a particular outcome of an uncertain quantity. For example, a particular realization of the uncertain quantity ${}_1x$ can be written as ${}_1x_i$.

The consideration of separate uncertain quantities is practical because the uncertainty in different parameters often has different sources and specific effects, and often different opportunities exist for reducing the uncertainty in the different quantities.

However, the separation of uncertainty into different quantities also necessitates the aggregation of uncertainty from these constituent parts back into an overall uncertainty in order to consider things such as expected consequences or expected utilities. This need to perform such calculations and to propagate uncertainty through engineering models and decisions motivates much of this dissertation.

Associated with each ${}_i x$ is a true probability distribution $p_i({}_i x)$. In general this true distribution is not known precisely, and it is more general to define $p_{ij}({}_i x) \in P_i(I)$ where $P_i(I)$ is the set of probability distributions for ${}_i x$ that is consistent with the available information.

3.6.3.4 Separating probabilistic and imprecise parameters for comparisons to existing methods

The uncertain quantities $\{{}_i x\}$ are a more general class than random variables because they can represent both probabilistic and imprecise quantities. While a probability bounds analysis can usually represent this combined information, other approaches cannot, so it is useful to make an additional restriction on the mathematical model that allows for comparison to traditional approaches.

As an example of a general uncertain parameter that is both probabilistic and imprecise, it may be that ${}_2 x$ is known to be normally distributed, but the mean of the distribution may not be known precisely; the mean is itself an uncertain quantity, ${}_1 x$ for example. If the variance is known to be σ^2, one can then write ${}_2 x \sim N({}_1 x, \sigma^2)$.

In some uncertainty models, such as probability bounds analysis, the DM could represent the uncertain quantity ${}_2 x$ directly. However, for comparison to formalisms that do not generalize both probability and interval theory, it is necessary to decompose the problem into purely probabilistic and purely imprecise parameters. For clarity, it is now useful to assume that the set of uncertain parameters is grouped, such that ${}_1 x$ through ${}_k x$ represent the imprecise parameters, and ${}_{k+1} x$ through ${}_n x$ represent probabilistic parameters. It is thus assumed that there are k imprecise parameters and $n-k$ purely probabilistic parameters. Associated with each of the $n-k$ probabilistic parameters $\{_i x \ \forall \ i = k+1,...,n\}$ is a precise distribution function $p_i(_i x)$ whose parameters may or may not be known. If $p_i(_i x)$ involved imprecise parameters, then really $p_i(_i x)$ represents a set of distributions $P_i(I)$ that are consistent with the available information. If the parameters of $p_i(_i x)$ are not known precisely, then these parameters must be expressed as uncertain, imprecise quantities $_i x$ (such that $i = 1,...,k$) in the model.

A necessary assumption of using this method is that the structure of the probabilistic uncertainties, that is, the forms of the probability distributions (such as Normal, Weibull, Gamma) are known precisely, but the parameters are not. The ability of probability bounds analysis to relax this assumption[10] is one of the advantages discussed in Chapter 5.

In general, the probabilistic parameters could be inherently random or based on subjective probabilities. The key distinction is how well the uncertainty is characterized. The determination of whether to represent the uncertainty using subjective probabilities or intervals is admittedly left to the discretion of the DM. However, one advantage of PBA methods is that they allow for both types of uncertainty to be represented in a well-defined manner, and an expert can even construct a p-box without explicitly making this decision. For example, the DM could construct a p-box by thinking directly about the set of distributions that he or she feels are consistent with the available information and his or her beliefs.

---

[10] A *general p-box* (defined in Section 4.1) relaxes this constraint. A *parameterized p-box* (defined in Section 4.1) maintains this constraint.

3.7 Summary

This chapter contains an overview of several uncertainty models, a step in answering the second motivating question posed in Section 1.5. The need for operational definitions was presented, and several uncertainty models were described. Probability theory and imprecise probability theory were discussed in detail, and a motivation for probabilities that are most generally imprecise and subjective was provided. Finally, the nature of an imprecise decision problem was captured by defining a general mathematical description of the problem. The next chapter describes probability bounds analysis (PBA) in more detail, and the following chapters discuss the application of PBA to design problems, including validating the answers to motivating questions 1 and 2 presented in Chapters 2 and 3 via example and general argument, and addressing motivating questions 3, 4, and 5.

CHAPTER 4:

PROBABILITY BOUNDS ANALYSIS (PBA)

In the previous chapter, several uncertainty models were explored, and it was suggested that engineers model uncertainty using probabilities that are most generally subjective and imprecise. It was also noted that computing with general imprecise probabilities is extremely expensive, relying on difficult mathematical programming problems. It is therefore useful to place additional constraints on the set of allowable distributions $P_i(I)$ for each uncertain parameter. One approach to further constraining the problem is called probability bounds analysis (PBA), in which the set $P_i(I)$ is limited to distributions that are contained in a probability box, or p-box.

P-boxes are introduced and defined in Section 4.1. The expressivity of the p-box uncertainty model is discussed in Section 4.2, particularly focusing on how p-boxes generalize traditional probability theory. Section 4.3 is an examination of methods for constructing p-boxes from available information. In Section 4.4, the interpretation of a p-box is considered. Methods for computing with p-boxes are presented in Section 4.5. Finally, Section 4.6 contains a preliminary description of decision-making with p-boxes, focusing on the calculation of the expected value of an uncertain parameter that is described by a p-box.

4.1 Probability boxes (p-boxes)

In probability bounds analysis (PBA), uncertainty is represented in a structure called a probability box, or p-box (Ferson and Donald 1998). A general p-box is defined as a set of cumulative probability distribution functions (CDFs) that is bounded from above and below. The CDF for a traditional random variable X is written as $F_X(x)$ and is defined in Equation (4.1). $F_X(x)$ represents the probability that the value of the uncertain quantity X is less than some set value x. In this dissertation, this notation is directly extended for all uncertainty quantities[11] X.

$$F_X(x) = P(X \leq x) \tag{4.1}$$

As defined by Bruns and Paredis (2006), the formal definition of the general p-box $\boxed{X}$ of some uncertain quantity X is a set of non-decreasing (cumulative probability) distribution functions (CDFs) constrained by the bounds in Equation (4.2).

$$\boxed{X} = \left\{ F_X(x) : \forall x \in \mathbb{R},\ \underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x) \right\} \tag{4.2}$$

In Equation (4.2), $\underline{F}_X(x)$ and $\overline{F}_X(x)$ are respectively the lower and upper cumulative probability bounds, and $F_X(x)$ is non-decreasing with x. Graphically, a p-box is a region bounded by two CDFs, as shown in Figure 4.1(a). Any distributions (such as those shown in Figure 4.1(b)) entirely inside the general p-box are considered consistent with the available information and possibly the true distribution. Any distribution that falls even partially outside of the p-box is considered inconsistent with the available information.

[11] The distinction is purely one of nomenclature. The use of the term "random variable" suggests inherent randomness. An uncertain quantity can reflect random behaviors, but also could be deterministic but unknown to the DM.

By definition, any CDF that does not violate the bounds in Equation (4.2) is acceptable and "contained" in the general p-box. As shown in Figure 4.2(a), these distributions could take any (non-decreasing) functional form. It is therefore useful to also define a parameterized p-box, such as shown in Figure 4.2(b), which limits the contained distributions to particular admissible types.

A parameterized p-box assumes a particular distribution type but leaves the parameters as imprecise quantities. For example, if uncertain parameter X is known to be normally distributed but with an imprecise mean, $\mu \in [\underline{\mu}, \overline{\mu}]$, and an imprecise

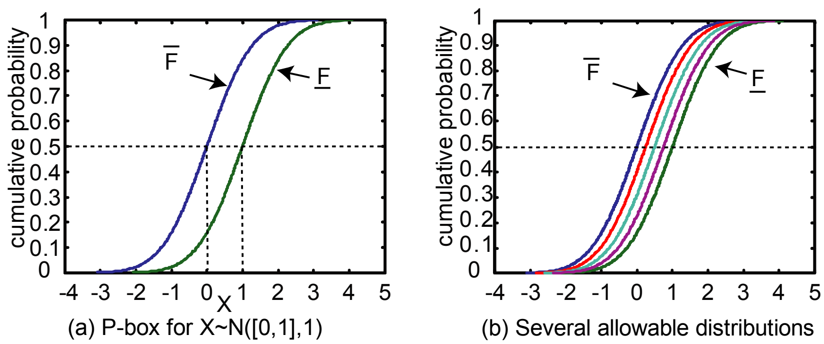


Figure 4.1. Example p-box

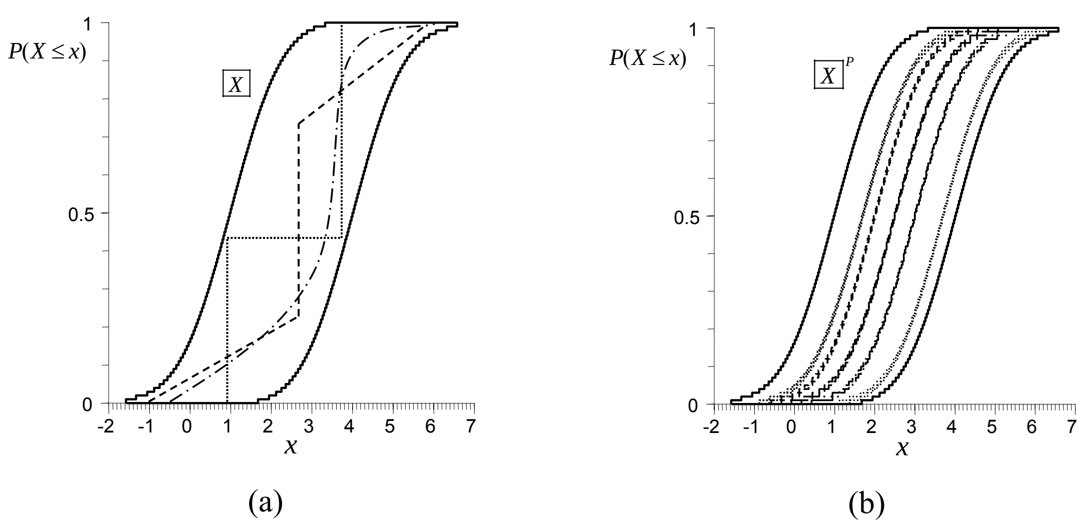


Figure 4.2. General and parameterized p-boxes with the same bounding functions but different admissible distribution examples.

standard deviation, $\sigma \in [\underline{\sigma}, \overline{\sigma}]$, then this total uncertainty could be written as $X \sim N([\underline{\mu}, \overline{\mu}], [\underline{\sigma}, \overline{\sigma}])$. This imprecise probability distribution corresponds to the parameterized p-box given in Equation (4.3), which is specific to a normal distribution.

$$\boxed{X}^P = \left\{ F_X(x; \mu, \sigma) = \Phi_{\mu,\sigma}(x) : \mu \in [\underline{\mu}, \overline{\mu}], \sigma \in [\underline{\sigma}, \overline{\sigma}] \right\} \tag{4.3}$$

In Equation (4.3), the superscript P denotes that the p-box is parameterized. The difference between general and parameterized p-boxes is displayed in Figure 4.2. Both of these p-boxes have the same bounds, in the sense that for all x_i, the cumulative probability $F_X(x_i)$ will have the same bounds, meaning that $\underline{F}_X(x_i) \leq F_X(x_i) \leq \overline{F}_X(x_i)$. However, the general p-box in Figure 4.2(a) contains an infinite number of non-decreasing functions that are not found in the parameterized p-box in Figure 4.2(b). A parameterized p-box will not contain all non-decreasing functions between its lower and upper bounding functions. The bounds themselves may not even be in the p-box, because the bounds on the cumulative probabilities are not guaranteed to have a functional form that is allowable.

For example, a parameterized p-box for a normal distribution with imprecise mean and variance is shown in Figure 4.3. The bounds of the p-box are not CDFs corresponding to normal distributions, as is readily concluded given the discontinuity in each bound at a cumulative probability of 0.5. This occurs because the bounds are formed by what are actually several allowable distributions, as shown in Figure 4.4. Recalling that the bounds on the mean are $\mu \in [\underline{\mu}, \overline{\mu}]$ and the bounds on the standard deviation are $\sigma \in [\underline{\sigma}, \overline{\sigma}]$, and accounting for the independence between the imprecision in the mean and standard deviation (meaning the smallest value of the mean could correspond to either the highest standard deviation or the lowest standard deviation), the actual bounds on the p-box correspond to the envelope of four distributions: $N(\underline{\mu}, \underline{\sigma})$, $N(\underline{\mu}, \overline{\sigma})$, $N(\overline{\mu}, \overline{\sigma})$, and $N(\overline{\mu}, \underline{\sigma})$. Any normal distribution with $\mu \notin [\underline{\mu}, \overline{\mu}]$ and/or $\sigma \notin [\underline{\sigma}, \overline{\sigma}]$ will have at least one point that falls outside the p-box shown in Figure 4.3.

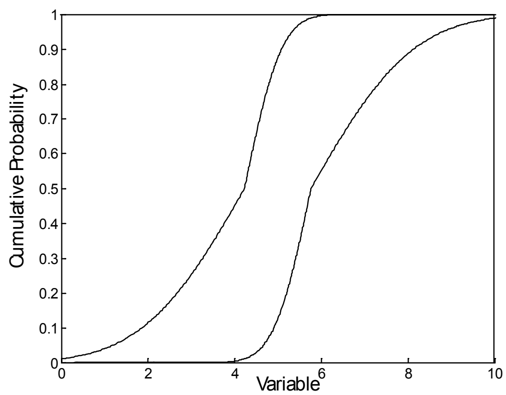


Figure 4.3. Resulting p-box

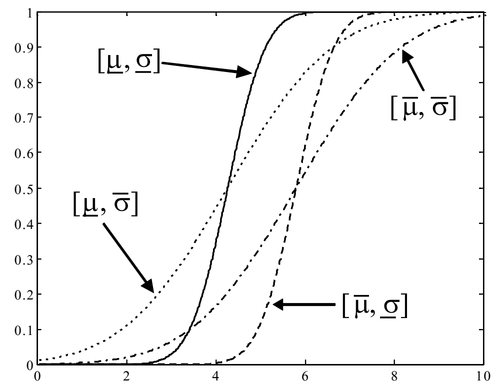


Figure 4.4. Forming bounds of the p-box

4.2 Expressivity of a p-box

A p-box is a more expressive generalization of both traditional probability distributions and interval representations, as is illustrated in Figure 4.5. The p-box incorporates both imprecision and probabilistic characterizations by expressing interval bounds on the cumulative probability distribution function (CDF) for a random variable. In this way, a p-box explicitly expresses both probability (represented by the shapes of the boundary CDFs) and imprecision (represented by the separation between the upper and lower bounds).

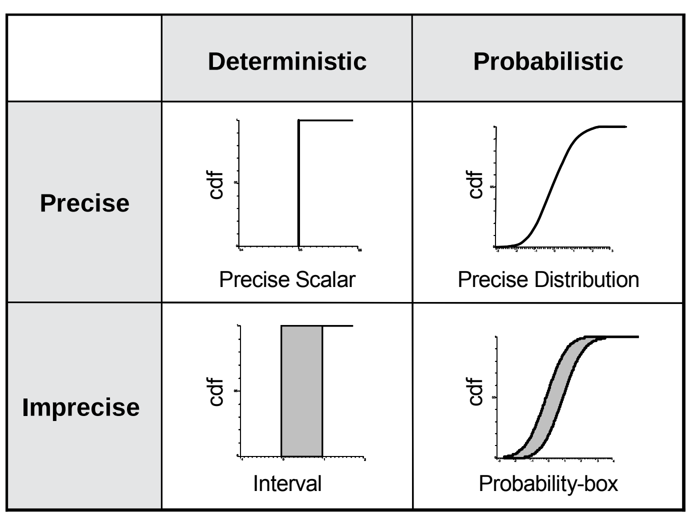


Figure 4.5. Dimensions of uncertainty

The p-box is general enough to represent intervals, scalars, and probability distributions, as well as imprecise probability distributions. An interval $X = [a,b]$ corresponds to the p-box defined by the probability bounds

$$\underline{F}_X(x) = \begin{cases} 0, & x < b \\ 1, & x \geq b \end{cases} \tag{4.4}$$

and

$$\overline{F}_X(x) = \begin{cases} 0, & x < a \\ 1, & x \geq a \end{cases} \tag{4.5}$$

A scalar a corresponds to the degenerate bounds shown in Equation (4.6).

$$\underline{F}_X(x) = \overline{F}_X(x) = \begin{cases} 0, & x < a \\ 1, & x \geq a \end{cases} \tag{4.6}$$

As an example of a precise probability distribution, a normally distributed random variable, $X \sim N(\mu,\sigma)$, corresponds to the p-box containing only one CDF, in which $\underline{F}_X(x) = \overline{F}_X(x) = \Phi_{\mu,\sigma}(x)$, where $\Phi_{\mu,\sigma}(x)$ is the cumulative distribution function of the normal distribution with mean μ and standard deviation σ.

The ability of p-boxes to reduce to traditional probability distributions is a major advantage of p-boxes over other uncertainty models such as imprecise probabilities. In cases in which a decision-maker (DM) has a large amount of information, a p-box approaches a precise probability distribution. For example, if $F_X^*(x)$ is the true, precise distribution, then $\underline{F}_X(x) \to F_X^*(x)$ and $\overline{F}_X(x) \to F_X^*(x)$ as the amount of information increases towards infinity.

A p-box is less expressive than a general imprecise probability because of the restrictions the type of distributions that can be considered consistent with the available information. In a general theory of imprecise probabilities, the only constraint placed on the set of distributions consistent with the available information is that the distributions really are consistent with the available information. PBA requires that the distributions be captured using the bounds shown in Equation (4.2) or as a parameterized p-box such as shown in Equation (4.3) for a normal distribution.

The bounding distributions of a p-box define intervals of (cumulative) probabilities. Walley (1991, 1996, 2000) has provided an example that shows how bounds on probabilities (e.g. p-boxes) cannot capture all possible states of imprecise information (the most concise description is Example 2 in (Walley 2000)). In Walley's example, the DM is asked to consider a football game with three possible outcomes for the home team, labeled as W (win), D (draw), and L (loss). The DM makes three qualitative judgments about his or her uncertainty in the outcome of the game:

i. 'not win' is at least as probable
ii. win is at least as probable as draw
iii. draw is at least as probable as loss

Judgment (i) can be represented in terms of upper and lower probabilities as either $\overline{P}(W) \leq \frac{1}{2}$ or $P(D \cup L) \geq \frac{1}{2}$. The other judgments cannot be expressed in terms of upper and lower probabilities. Walley notes that the statement $\underline{P}(W) \geq \overline{P}(D)$ is too strong for judgment (ii). The reason is that the intervals for $P(W)$ and $P(D)$ can actually overlap. For example, assume that (ii) is the only constraint. Then the following are possible intervals: $P(W) = [0,1]$ and $P(D) = [0,1]$. For any value in $P(W) = [0,1]$ it is possible to find at least one value in $P(D) = [0,1]$ such that "win is at least as probable as draw." However, the true value must satisfy the following relations: $P(W) + P(D) + P(L) = 1$ and $P(W) \geq P(D)$.

Rather than being expressible as ordered upper and lower probabilities, the three judgments should be seen as constraints on a coherent lower prevision of the form: (i) $P(W) \leq \frac{1}{2}$; (ii) $P(W) \geq P(D)$; and (iii) $P(D) \geq P(L)$. Naturally, the constraint $P(W) + P(D) + P(L) = 1$ also exists. This set of constraints defines a closed convex polyhedron, which in general cannot be defined by independent bounds on the individual events.

4.3 Constructing p-boxes

There are several ways to construct p-boxes (Ferson, et al. 2002b, Ferson, et al. 2005), depending on the type of information available. In this dissertation, a procedure is presented that constructs p-boxes from statistical data samples using 95% confidence intervals on the parameters of a known distribution type. While the distribution type will not always be known, in engineering applications it is common that some theoretical knowledge can guide the selection of a distribution type (Utkin and Augustin 2003). However, such knowledge is not a requirement for using PBA. Probability boxes also can be constructed based on distribution-free methods, such as using just statistical data or moments (Ferson 2002) or by using the empirical distribution and the Kolmogorov-Smirnov statistic to form bounds on the true distribution (Ferson, et al. 2005). If the engineer believes that several distributions are consistent with the available information, he or she could assume different types of distribution, construct p-boxes for each, and then take the envelope of all of these to form the most general p-box.

As noted above, the method of constructing a p-box from statistical data developed and used in this thesis assumes that the underlying probability distribution type for an uncertain quantity is known. The p-box is constructed in order to reflect the DM's lack of information (imprecision) about the parameters of this distribution. The method is developed and demonstrated with the assumption that the true distribution is a normal distribution (i.e. Gaussian). The generalization of this method to other distributions is addressed in Section 4.3.3.

4.3.1 Constructing p-boxes for normally distributed uncertain parameters

In this example, the DM assumes the uncertain quantity X is normally distributed, but with unknown mean and standard deviation:

$$X \sim Normal(\mu, \sigma) \tag{4.7}$$

The DM estimates the true but unknown μ and σ using the unbiased point estimates shown in the following equations, in which the x_i's are the sample observations, and n is the sample size.

$$\hat{\mu} = \overline{x} = \tfrac{1}{n}\sum_{i=1}^{n} x_i \tag{4.8}$$

$$\hat{\sigma}^2 = s^2 = \tfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{4.9}$$

These quantities are called respectively the sample mean (4.8) and sample variance (4.9) and are commonly used in pure probabilistic approaches. In order to construct a p-box, the point estimates of the parameters are broadened into confidence intervals. In this development, a 95% confidence level is used, but any confidence level is allowable, as noted in the next section.

Since the set $\{x_1, x_2, ..., x_n\}$ is a random sample from a normal distribution, the sampling distribution of the statistic shown in Equation (4.10) is the t distribution with $n-1$ degrees of freedom.

$$t = \frac{\hat{\mu} - \mu}{s/\sqrt{n}} \tag{4.10}$$

Letting $t_{\alpha/2,n-1}$ be the upper $\alpha/2$ percentile of the t distribution with $n-1$ degrees of freedom, Equation (4.11) can be found.

$$P\left\{-t_{\alpha/2,n-1} \le t \le t_{\alpha/2,n-1}\right\} = 1-\alpha \tag{4.11}$$

Substituting for t from Equation (4.10) into Equation (4.11) and solving for the mean μ, t a $(1-\alpha)100\%$ confidence interval for the mean is found, given in the following equation.

$$[\underline{\mu}, \overline{\mu}] = [\hat{\mu} - t_{\alpha/2,n-1}\, s/\sqrt{n}\,, \hat{\mu} + t_{\alpha/2,n-1}\, s/\sqrt{n}] \tag{4.12}$$

Since $\{x_1, x_2, ..., x_n\}$ is a random sample from a normal distribution, it can be shown that the sampling distribution for the variance is chi-square with $n-1$ degrees of freedom, as shown in Equation (4.13), where n is the sample size and s^2 is the sample variance (Hines, et al. 2003).

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \tag{4.13}$$

To develop the confidence interval, Equation (4.14) is noted.

$$P\left\{\chi^2_{1-\alpha/2,n-1} \le \chi^2 \le \chi^2_{\alpha/2,n-1}\right\} = 1-\alpha \tag{4.14}$$

Substituting for χ^2 from Equation (4.13) into Equation (4.14) and solving for the variance σ^2, one arrives at a $(1-\alpha)100\%$ confidence interval for the variance, given in Equation (4.15).

$$\left[\underline{\sigma}^2, \overline{\sigma}^2\right] = \left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right] \tag{4.15}$$

A table of t and χ^2 values is found in most probability and statistic books, such as (Hines, et al. 2003).

4.3.2 Choosing the confidence level for p-box construction

In general, there is no rule for selecting the confidence level at which to construct a p-box. However, a confidence level has a clear interpretation; if something, such as confidence interval on a point estimate, is constructed at the 95% confidence level, it means that if the experiment were repeated many times and a 95% confidence interval constructed for each repetition, then 95% of those intervals would contain the true value. An individual DM must understand the consequences of confidence levels and choose one that is appropriate for the problem and preferences at hand.

In this way, the selection of a confidence level has some of the same limitations as the selection of a prior in Bayesian analysis. However, there are major advantages to using confidence levels and p-boxes instead of Bayesian analysis. The first is the separation of imprecision and irreducible uncertainty, as discussed in Section 3.3.5. The second is that confidence levels have a clear meaning and are consistent, in that (for a given uncertain quantity) the 99% confidence interval always contains the 95% confidence interval which always contains the 90% confidence level and so on. Conversely, the use of different prior (i.e. input) distribution in Bayesian analysis can lead to very different posterior (i.e. output) distributions, yet it is not obvious in a posterior distribution what the initial assumptions were.

4.3.3 Constructing p-boxes for other distributions

In Section 4.3.1, the construction of p-boxes was illustrated under the assumption that the true distribution is normal. In order to construct p-boxes for other distributions using this method, it is only necessary to find the appropriate statistical descriptions for confidence intervals on the distribution's parameters. For example, if the true distribution is assumed to be a Weibull distribution, then one needs to create confidence intervals for the parameters of the Weibull distribution. More specifically, if Z follows a Weibull distribution, then $Z \sim Weibull(\alpha, \beta)$, and the parameters of concern are α and β. These parameters are not easily defined in terms of sample statistics, and their relationships to the mean and variance are given in Equations (4.16) and (4.17).

$$\mu = \beta \cdot \Gamma(1-\alpha^{-1}) \tag{4.16}$$

$$\sigma^2 = \beta^2 \cdot \Gamma(1-2\alpha^{-1}) - \Gamma^2(1-\alpha^{-1}) \tag{4.17}$$

One way of estimating α and β is using maximum likelihood (or log-likelihood) estimates. Confidence intervals can then be constructed using bootstrapping methods (Efron and Tibshirani 1993). These methods are not explicitly tested in this dissertation. This section discusses how it is possible in principle to construct such p-boxes, and the dissertation does contain an example (0) that constructs a p-box from sample data assuming the uncertain parameter is normally distributed. The actual use of maximum likelihood estimates and bootstrapping techniques to create p-boxes for non-normal uncertain parameters is left for future work.

4.4 Interpreting a p-box

For illustration, it is easiest to consider a p-box as constructed at the 100% confidence level, though in most practical problems such a p-box would be infinite. In this case, the p-box expresses the range of all CDFs that are still deemed possible based on existing information. For example, assume that for all practical purposes, X is a random variable, and engineers have strong theoretical information that X is normally distributed with known variance $\sigma^2=1$. However, the engineers can only characterize the mean imprecisely, bounding it in the interval $\mu=[0,1]$. Extending the notation of probability, one can write

$$X \sim N(\mu,\sigma^2)=N([0,1], 1). \tag{4.18}$$

The corresponding p-box is shown on page 118 in Figure 4.1. In this case, the bounds on the p-box are defined by the two distributions, $\overline{F} \sim N(0,1)$ and $\underline{F} \sim N(1,1)$. The true CDF is unknown, and any of the infinite number of normal CDFs inside the p-box with variance of one could be the true distribution. However, any distribution that falls partially or entirely outside of the p-box is inconsistent with the present state of information.

Vertical slices of the p-box yield intervals on the cumulative probability for a particular realization. For example, a vertical slice at zero yields the interval for the cumulative probability of [0.1587, 0.5]. This means that the probability that X is less than zero is between 0.1587 and 0.5, but one does not have enough information to specify a precise probability within that interval. Horizontal slices of a p-box result in intervals on the quantiles of the cumulative probability. For example, a slice at the median (cumulative probability =0.5) gives the interval [0,1] for the median.

4.5 Computing with p-boxes

Although not quite as expressive as imprecise probabilities, p-boxes have the advantage that relatively efficient algorithms have been developed to compute with p-box uncertainties. For examples of p-box propagation algorithms in the literature, see the work of Williamson and Downs (1990), Ferson and co-authors (Ferson and Ginzburg 1996, Ferson and Donald 1998, Oberkampf, et al. 2002a, Ferson and Hajagos 2004), and Berleant and coauthors (Berleant 1993, Berleant and Goodman-Strauss 1998, Berleant, et al. 2003, Berleant and Zhang 2004). The method of Ferson and co-authors is based on the work of Williamson and Downs and uses interval-based operations to propagate uncertainty. Berleant's method uses optimization rather than interval-based operations. Together, these methods can be called *dependency bounds convolution* (DBC) methods. This method is taken as the base case in this dissertation. Several recent methods for computing with p-boxes are explored and developed by Bruns, Paredis, and Ferson (2006). These methods are mentioned briefly in Section 4.5.2. Much of this section is adapted from this paper and work included in the master's thesis of Bruns (2006).

4.5.1 Dependency Bounds Convolution (DBC)

Dependency bounds convolution (or DBC) (Ferson and Donald 1998, Ferson 2000, Tucker and Ferson 2003, Ferson and Hajagos 2004) is a term used to describe a class of rigorous methods for propagating p-boxes through mathematical models. The results of DBC are rigorous in the sense that the resultant probability bounds are guaranteed to contain the true probability distribution of the uncertain quantity for any possible dependence relationship between the inputs—assuming that the input p-boxes were themselves rigorous (i.e. contain the true distributions). These probability bounds can also be described as best-possible in the sense that they are as tight as possible given the information provided in the input p-boxes, assuming there are no repeated variables (see Section 4.5.2). Finally, DBC calculations are applicable towards non-parameterized p-boxes. This means that no assumptions are made about the true probability distribution other than that it is contained within the p-boxes bounding functions.

Two methods for DBC were developed independently by Williamson and Downs (Williamson 1989, Williamson and Downs 1990) and Berleant and co-authors (Berleant 1993, Berleant and Goodman-Strauss 1998, Berleant and Zhang 2004). The method of Berleant is also referred to as Distribution Envelope Determination (or DEnv), but in this dissertation the two methods are grouped under the name DBC. Regan, Ferson, and Berleant have shown that the method of Williamson and Downs and DEnv are equivalent for binary operations of p-boxes defined on the positive real numbers (Regan, et al. 2004).

Although it is unnecessary to fully describe mathematical details of the methods for DBC that are developed elsewhere (Williamson 1989, Williamson and Downs 1990, Ferson and Donald 1998), it is helpful to sketch in outline how these methods function. The DBC calculation begins with a bounding discretization of the input p-boxes. This is done by partitioning the p-box into a set of n horizontal slices. Each slice is fully described by a probability mass (the vertical height of the slice) and an interval corresponding to lower and upper bounds on a subset of the domain of the uncertain quantity. For example, the p-box $\boxed{X}$ shown in Figure 4.6 is discretized into four slices, each of probability mass 0.25. The discretized p-box contains the true p-box. The second slice from the bottom is associated with the closed interval $\left[\underline{x}_2, \overline{x}_2\right]$.

The algorithms of Williamson and Downs (1990) provide deterministic and rigorous approximations of binary functions of discretized p-boxes. Any binary function of p-boxes can be analyzed in the context of a Cartesian product of the input p-boxes. This utilization of a Cartesian product was first proposed by Yager (1986) for the convolution of Dempster-Shafer structures. Consider some binary function of uncertain quantities, $\boxed{Z} = f\left(\boxed{X},\boxed{Y}\right)$. If the uncertain inputs, $\boxed{X}$ and $\boxed{Y}$, are discretized into m and n slices, respectively, then the resultant Cartesian product is an mn-element list of interval-mass pairs.

Suppose that the discretization slices for the two inputs are evenly distributed along the cumulative probability axis. Then every slice for $\boxed{X}$ has probability mass $1/m$, and every slice for $\boxed{Y}$ has probability mass $1/n$. Also, the slices of $\boxed{X}$ are labeled as the intervals $x_i = \left[\underline{x}_i, \overline{x}_i\right], i = 1,...,m$ and the slices of $\boxed{Y}$ as $y_j = \left[\underline{y}_j, \overline{y}_j\right], j = 1,...,n$. Then the ijth-element of the Cartesian product for $\boxed{Z} = f\left(\boxed{X},\boxed{Y}\right)$, assuming statistical independence between the inputs, is $\left(f(x_i, y_j), \frac{1}{m} \times \frac{1}{n}\right)$ where $f(x_i, y_j)$ is the interval extension of the intervals x_i and y_j. The resultant p-box, $\boxed{Z}$, in the case of independence, is then an ordered stacking of the slices $\left(f(x_i, y_j), \frac{1}{m} \times \frac{1}{n}\right)$.

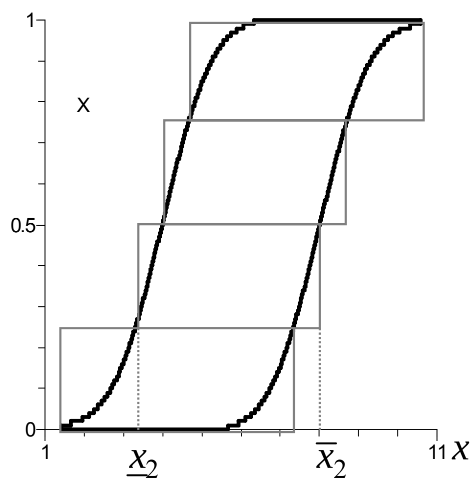


Figure 4.6. A discretized p-box.

Although this result for the statistically independent case is useful, a more powerful quality of DBC methods is the ability to determine probability bounds in the case of unknown dependence between the inputs. Although engineers often lack knowledge about the true dependence between uncertain quantities, they still tend to assume independence in their models—an assumption that likely results in incorrect conclusions. Both DEnv and DBC provides methods for computing without any assumptions of independence or dependence. Essentially, a p-box can be computed that covers all possible dependency scenarios.

4.5.2 Limitations of DBC methods

DBC methods are rigorous and efficient, but they must overcome at least two obstacles before they can be used effectively in engineering design. First, both of these approaches depend strongly on the methods of interval arithmetic for which the presence of repeated variables can result in over-conservative (i.e. not best-possible) solution bounds.

In general, DBC are rigorous but not best possible. In interval arithmetic (and by extension in probability bounds analysis), upper and lower bounds are best possible if the upper bound is as low as possible and the lower bound is as high as possible without conflicting with the true state of uncertainty. Bounds are rigorous as long of the best possible bounds are included between them.

Repeated variables in an expression often lead to bounds that are not best possible with interval arithmetic and DBC. This can be summarized by the failure of the distributive law (Ferson 2002, Muhanna and Mullen 2004), which means, among other things, that in interval arithmetic it is not always true that $A \cdot B + A \cdot C = A \cdot (B + C)$ (Ferson 2002). This is a definite limitation of interval methods and PBA. However, it is very useful that the methods are rigorous, meaning that the true interval $A \cdot (B + C)$ is always contained in the calculated interval $A \cdot B + A \cdot C$, so the calculated bounds are true, but not best possible. Although overly conservative results can be avoided using sub-interval reconstitution methods (Moore 1979, Ferson and Hajagos 2004, Ferson, et al. 2004b), such methods are prohibitively expensive in realistic engineering problems with a large number of imprecise quantities.

A second limitation of DBC is that it is not application to engineering problems that involve black-box models. For this dissertation, a black-box model is defined as any model that is not an easily accessible, closed form algebraic equation. For example, a simulation is a black-box from the perspective of analysis, as the interactions between the uncertainty analysis and simulation can only be through input and output behavior; an uncertainty analysis tool cannot interact with the inside (e.g. code) of the simulation. It may be possible to recode these models using languages and methods that are suitable for p-box computations, but this would be prohibitively costly in most industry problems, where often legacy code is the rule rather than the exception.

If a model is algebraic, it is easily translated into a white-box such that a p-box-based or other uncertainty analysis method can interact directly with the core content of the model. Some black-box propagation methods for interval propagation have been developed. Trejo and Kreinovich (2001) and Kreinovich and Ferson (2004) have developed a randomized algorithm for propagating interval uncertainty through black-box models, but the method assumes that the black-box model is broadly linear in the region of sampling.

Bruns and co-authors (Bruns 2006, Bruns and Paredis 2006, Bruns, et al. 2006) have summarized other methods and developed a new method called p-box convolution sampling (PCS). These methods fall roughly into two classes: methods for parameterized p-boxes and methods for general p-boxes. The four methods compared are DBC, PCS, optimized parameter sampling (OPS), and double loop sampling (DLS). The three newer methods (PCS, OPS, and DLS) all involve sampling of the distributions in some way. As a result, the rigor of DBC (in terms of the true distribution being guaranteed to be inside the resultant bounds) is lost. However, the newer methods are all applicable to black-box models.

OPS and DLS apply only to parameterized p-boxes. Significant savings are possible when dealing with parameterized p-boxes because the function form of the CDFs is well defined, so the sampling in some sense is a sampling of the parameters of these distributions. The general p-box case is more complicated because there are an infinite number of distribution types included in the p-box. This is value of the PCS method; it extends the application of PBA to black-box models when the uncertainty is described using a general p-box.

The ability of PBA to be used with black-boxes is significant. Bruns and co-authors have shown (Bruns 2006, Bruns and Paredis 2006, Bruns, et al. 2006) reasonable performance of these methods in relatively simple engineering examples. For this dissertation, the existence of methods for propagating p-boxes through black-box models is taken as sufficient to warrant the continued study of the benefits of using p-boxes in engineering design. Once the potential benefit is established, future work can re-examine the computational costs and issues associated with applying the methods to more complex problems. The demonstration of potential benefit is necessary to warrant additional research into improved computational methods; unless the potential benefit is identified, there is no reason to expend resources researching the computational methods.

4.6 P-boxes and decision making

P-boxes are a set of probability distributions describing some quantity. Using the methods described in Section 4.5, it is possible to propagate uncertainty and p-boxes through calculations until a p-box for some end objective results. For example, a p-box for the expected utility of a design alternative can be constructed.

Such a p-box for utility would express the set of CDFs that could possibly describe the utility of a decision alternative given the available evidence. In this dissertation, it is assumed that the utility functions are known precisely (i.e. can be fully elicited). In the context of the general model described in Section 3.6, this means that the set $U(I)$ contains only one function (recall that I represents information available to the DM).

In general, the utility function will not be known precisely and the set $U(I)$ will contain multiple distributions. However, the focus is currently on the uncertainties in probability distributions or the set $P(I)$, so the more general and complicated case of imprecise probability distributions and imprecise utility functions is deferred for future work. Like many things, engineering design uncertainty must be addressed one piece at a time, with a hopeful unifying theory emerging in the future from the components.

4.6.1 Intervals of expected utility

Utility theory defines the optimal decision as the selection of the alternative with the greatest expected utility. When uncertainty is represented using a p-box, there is in general no single expected utility associated with an alternative. This can be seen by revisiting the definition of mathematical expectation of a random quantity X. In traditional statistics, the expected value is defined as the following:

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x \cdot p(x) \cdot dx \qquad \text{(continuous random variable)} \\ \mathrm{E[X]} &= \sum_{i=1}^{\infty} x_i \cdot p(x_i) \qquad \text{(discrete random variable)} \end{aligned} \tag{4.19}$$

Considering just the continuous case for now, if there are multiple PDFs $p_j(x)$, then there are multiple expectations, such as defined in Equation (4.20).

$$E_j[X] = \int_{-\infty}^{\infty} x \cdot p_j(x) \cdot dx \text{ for all } j \text{ such that } p_j(x) \in P(I) \tag{4.20}$$

When the set of allowable distributions $P(I)$ is those distributions for which the CDF falls entirely inside the p-box of the uncertain quantity X, then there are an infinite number of $p_j(x)$, and consequently an infinite number of $E_j[X]$. However, this set of expected values forms an interval, given by Equation (4.21). Using Equations (4.20) and (4.21), the notion of expected value is extended from a single random distribution to a more general uncertain quantity.

$$
\begin{aligned}
E[X] &= \left[\underline{E}[X], \overline{E}[X]\right] \\
\underline{E}[X] &\equiv \min_j (E_j[X]) \\
\overline{E}[X] &\equiv \max_j (E_j[X])
\end{aligned}
\tag{4.21}
$$

4.6.2 Calculating the expectation of a p-box

The interval in Equation (4.21) can be found easily, as the it can be shown that the bounds on the expected value are found using the bounding distributions of the p-box. Before developing the proof, the mathematical notations from the general model and for the p-box definitions of Section 4.1 are revisited. Recalling Equation (4.2) a p-box $\boxed{X}$ for uncertain quantity X is defined in terms of bounding CDFs, namely $\underline{F}_X(x)$ and $\overline{F}_X(x)$. Now, two particular PDFs $\underline{f_X}(x)$ and $\overline{f_X}(x)$ are defined in Equations (4.22) and (4.23).

$$
\underline{f_X}(x) \in P(I), \text{ such that } \underline{F}_X(x) = \int_{-\infty}^{x} \underline{f_X}(t) \cdot dt \tag{4.22}
$$

$$
\overline{f_X}(x) \in P(I), \text{ such that } \overline{F}_X(x) = \int_{-\infty}^{x} \overline{f_X}(t) \cdot dt \tag{4.23}
$$

With these distributions defined, the bounds on expected value can be re-defined in terms of the bounding distributions as in Equation (4.24). A proof follows.

$$
\begin{aligned}
E[X] &= \left[\underline{E}[X], \overline{E}[X]\right] \\
\underline{E}[X] &\equiv \min_j (E_j[X]) = \int_{-\infty}^{\infty} x \cdot \overline{f_X}(x) \cdot dx \\
\overline{E}[X] &\equiv \max_j (E_j[X]) = \int_{-\infty}^{\infty} x \cdot \underline{f_X}(x) \cdot dx
\end{aligned}
\tag{4.24}
$$

Equation (4.24) is proved using reasoning based on first-order stochastic dominance and the definition of a p-box. The proof of first order dominance is based on that by Levy (1998), but the extension to the p-box problem is a new contribution[12]. The proof begins with a few definitions. Let $F(x)$ and $G(x)$ represent two cumulative distribution functions, and let $f(x)$ and $g(x)$ represent the corresponding probability mass functions, such that $f(x) = F'(x)$ and $g(x) = G'(x)$. The mathematical expectation is defined with a subscript indicating the distribution with which the expectation is to be calculated. For example, assume some function $h(x)$. Then define $E_F[h(x)] = \int f(x)h(x)dx$ and $E_G[h(x)] = \int g(x)h(x)dx$.

A set N_1 is defined as the set of non-decreasing functions, such that for all $N(x) \in N_1$, $N'(x) \geq 0$ for all x in the domain of $N(x)$ (the derivative with respect to x is non-negative). The condition of first degree stochastic dominance can be summarized in Equation (4.25), which holds for all x and all $N(x) \in \mathrm{N}_1$, with strict inequality for at least one x_0 and at least one $N_0(x) \in N_1$.

$$F(x) \leq G(x) \quad \Leftrightarrow \quad E_F[N(x)] \geq E_G[N(x)] \tag{4.25}$$

The proof is conducted with the additional assumption that x is bounded from above and below. This means that there exists some a and b, $b \geq a$ for which $a \leq x \leq b$. An extension of the proof of Equation (4.25) to unbounded random variables is given in Hanoch and Levy (1969) and Tesfatsion (1976). The following proof is made in two parts. First, the consequences of Equation (4.25) with respect to calculating the expected value of a p-box are demonstrated. Second, Equation (4.25) is proved.

---

[12] Huber (1981) and Walley (1981) have proved similar relationships for lower previsions that are 2-monotone.

4.6.2.1 Bounds on expected value correspond to bounds on a p-box

Consider what Equation (4.25) means for the distributions in Figure 4.7, in which clearly $\underline{F}(x) \le \overline{F}(x)$, a relationship which is also guaranteed by the definition of a p-box in Equation (4.2). Then by Equation (4.25), $E_{\underline{F}}[N(x)] \ge E_{\overline{F}}[N(x)]$ for any $N(x) \in N_1$. Expanding with the definition of the mathematical expectation results in Equations (4.26) and (4.27).

$$E_{\underline{F}}[N(x)] = \int_a^b \underline{f}(x) \cdot N(x) \cdot dx \tag{4.26}$$

$$E_{\overline{F}}[N(x)] = \int_a^b \overline{f}(x) \cdot N(x) \cdot dx \tag{4.27}$$

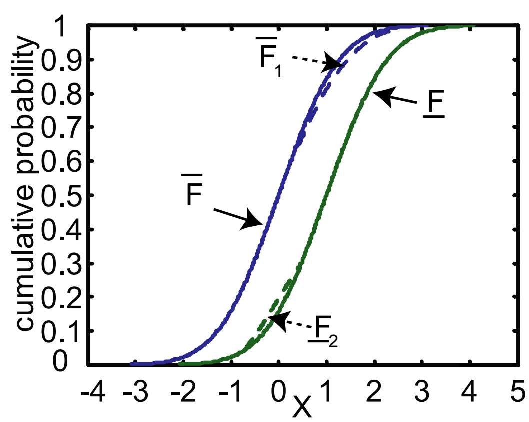


Figure 4.7. Calculating expected value of a p-box

The goal of this proof is to show bounds on the expectation of the actual random variable X. In this case, $N(x) = x$, which is a non-decreasing function; that is, $x \in N_1$. This can be seen by comparing Equations (4.26) and (4.27) respectively with Equations (4.28) and (4.29).

$$E_{\underline{F}}[X] = \int_a^b \underline{f}(x) \cdot x \cdot dx \tag{4.28}$$

$$E_{\overline{F}}[X] = \int_a^b \overline{f}(x) \cdot x \cdot dx \qquad (4.29)$$

Combining Equations (4.26)-(4.29) with the relationship that $\underline{F}(x) \le \overline{F}(x)$, it must be that $E_{\underline{F}}[X] \ge E_{\overline{F}}[X]$. This result orders the expected value of the two bounding distributions. What about the distributions inside the p-box? By definition in Equation (4.2), any distribution inside the p-box, such as F_i, must obey the relationship $\underline{F}(x) \le F_i(x) \le \overline{F}(x)$. Based on this relationship, $F_i(x)$ stochastically dominates $\overline{F}(x)$, and consequently $E_{F_i}[X] \ge E_{\overline{F}}[X]$. In other words, the expected value of X calculated using the upper-bounding cumulative distribution $\overline{F}(x)$ must be less than or equal to the expected value calculated using any distribution inside the p-box. As such, $E_{\overline{F}}[X]$ must be the lower bound on the expected value of X, thus proving the second line of Equation (4.21). Notice this holds for any distribution, even one such as $F_1(x)$ in Figure 4.7 that is mostly coincident with $\overline{F}(x)$, but for which the relationship $\overline{F}(x) \ge F_1(x)$ holds for all x and $\overline{F}(x_0) > F_1(x_0)$ for some x_0.

Since it is also true by definition in Equation (4.2) that $\underline{F}(x) \le F_i(x)$, it is also true that $\underline{F}(x)$ stochastically dominates $F_i(x)$, which implies that $E_{\underline{F}}[X] \ge E_{F_i}[X]$. In other words, the expected value of X calculated using the lower bounding distribution of $\underline{F}(x)$ is greater than or equal to the expected value calculated using any distribution inside the p-box. As such, $E_{\underline{F}}X$ must be the upper bound on the expected value of X, thus proving the third line of Equation (4.21). Notice that this relationship holds for any $F_i(x)$ inside the p-box, included a distribution such as $F_2(x)$ in Figure 4.7 that is mostly coincident with $\underline{F}(x)$ but for which the inequality $\underline{F}(x) \le F_2(x)$ holds for all x and $\underline{F}(x_0) < F_2(x_0)$ for some x_0.

A slight refinement of the proof would be required to change the inequalities of the equations into strict inequalities. However, this is not necessary for the current proof. The goal was to show that the lower (upper) bound on expected value can be calculated using the upper (lower) bounds on the p-box. In order to do this, it is not necessary to show that those are uniquely the lowest (greatest), but only that there are no distributions lower (greater). Based on the proof given, there could be other distributions inside the p-box that yield the same values as the bounding distributions. In practice, this cannot happen for $N(x) = x$, but a more detailed proof is unnecessary because the claim that the bounding distributions yield bounds on the expected value holds with the less strict proof.

In summary, the condition of first degree stochastic dominance shown in Equation (4.25) leads directly to a proof of Equation (4.21). For completeness, a summarized proof of Equation (4.25) is now provided.

4.6.2.2 Proof of first degree stochastic dominance relationship

In the previous proof, the implication in Equation (4.25) is only used in one direction, as shown in Equation (4.30). As such, only this relationship is proved.

$$F(x) \le G(x) \quad \Rightarrow \quad E_F[N(x)] \ge E_G[N(x)] \tag{4.30}$$

The proof, based on that by Levy (1998), begins by noting that if $I_1(x) \equiv G(x) - F(x)$, then $F(x) \le G(x)$ for all x implies that $I_1(x) \ge 0$ for all x, as given in Equation (4.31).

$$I_1(x) \equiv G(x) - F(x) \ge 0 \tag{4.31}$$

Similarly, the consequence $E_F[N(x)] \ge E_G[N(x)]$ implies that $\Delta \equiv E_F[N(x)] - E_G[N(x)] \ge 0$. The quantity Δ can be expanded as in Equation (4.32), which can be rewritten as Equation (4.33).

$$\Delta \equiv E_F[N(x)] - E_G[N(x)] = \int_a^b f(x)N(x)dx - \int_a^b g(x)N(x)dx \tag{4.32}$$

$$\Delta = \int_a^b (f(x) - g(x))N(x)dx \tag{4.33}$$

This integral is now evaluated using integration by parts. Recall that integration by parts evaluates an integral according to Equation (4.34).

$$\int u dv = uv - \int v du \tag{4.34}$$

Letting $u = N(x)$ and $dv = \left(f(x) - g(x) \right) dx$, one finds $du = N'(x)$ and $v = \int (f(x) - g(x))dx$. Recalling that $F(x) = \int f(x)dx$ and $G(x) = \int g(x)dx$, $v = F(x) - G(x)$. Thus, Equation (4.33) can be rewritten as Equation (4.35).

$$\Delta = N(x) \cdot [F(x) - G(x)]\Big|_a^b - \int_a^b N'(x)[F(x) - G(x)]dx \tag{4.35}$$

Based on the assumption that x is bound from below by a and above by b, $F(a) = G(a) = 0$ and $F(b) = G(b) = 1$. Consequently, $N(x) \cdot [F(x) - G(x)]\Big|_a^b = 0$. This reduces Equation (4.35) to Equation (4.36) by removing the first term and distributing the negative sign.

$$\Delta = \int_a^b N'(x)[G(x) - F(x)]dx \tag{4.36}$$

Recalling Equation (4.31), $G(x) - F(x) \geq 0$. Also, $N'(x) \geq 0$ since $N(x) \in N_1$. Since the integral of any non-negative number is non-negative, it must be that $\Delta \geq 0$, thus proving that $E_F[N(x)] \geq E_G[N(x)]$, the required consequence of Equation (4.30).

4.6.3 Decision making with intervals of expected utility

In this section, the focus is on one important consequence of imprecision and the existence of interval of expected utility—namely, that it can result in *indeterminacy in decision making*. Indeterminacy means that based on the available information, one cannot determine which decision alternative is most preferred. A complete discussion of the issue of resolving indeterminacy is left for Chapter 8. This section merely introduces the problem and briefly explains some policies.

In general, there are three possible scenarios of preference between alternatives A and B: A is preferred to B, B is preferred to A, or the DM is indifferent between A and B. Note that indifference implies a strict equality in preference, meaning the DM would willingly trade A and B. When utilities are used to reflect preference, these relationships can be determined by the inequality or equality of the expected utilities (von Neumann and Morgenstern 1944). However, when imprecision exists, the expected utilities become intervals (as described in Section 4.6.1), and such comparisons become more complicated.

For example, consider the intervals of expected utility for two alternatives (A and B) shown in Figure 4.8(a). Because the intervals do not overlap, a clear choice can be made. Specifically, alternative A is clearly the better. In the example in Figure 4.8(b), the intervals overlap. Since the true expected utility of B can lie anywhere in the given interval, the point labeled b_1 is possible. Similarly, both a_1 and a_2 are possible true values for the expected utility of A. Notice that a_1 is greater than b_1, but a_2 is less than b_1. Consequently, the available evidence is *indeterminate*; the DM cannot determine which alternative is the most preferred, nor can the DM determine that he or she is strictly indifferent. In order to make elimination decisions in the presence of imprecision, different methods are needed.

In some cases, a rigorous reduction of the intervals can be performed by

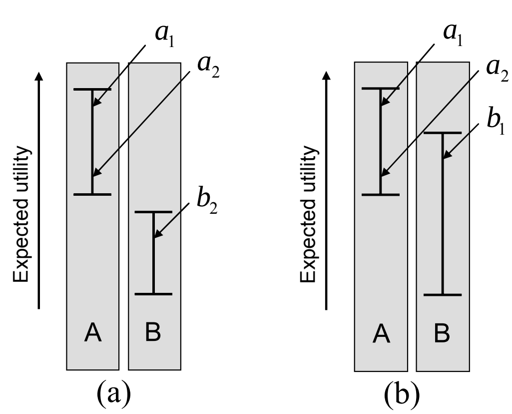


Figure 4.8. Intervals of expected utility

considering additional information that is not directly captured in the representation. Such methods are discussed in more detail in Chapter 8. In other cases, new information can be collected, a process addressed in Chapter 9. However, in some cases, a decision maker may elect to make an arbitrary choice. *Arbitrary* in this sense does not necessarily imply *without guidance* or *random*. Several policies are possible to guide arbitrary choice, including Γ-maximin (Berger 1985) and the Hurwicz-criterion (Arrow and Hurwicz 1972). A Γ-maximin policy says that given indeterminacy in a maximization problem, a DM should select the alternative with the highest lower-bound. This is a conservative policy in that it seeks to mitigate the worst-case.

The policies are referred to as arbitrary because the choices they lead to are not determined by a strict normative, rational theory such as utility theory. The overlapping intervals of expected utility actually imply that based on the available information, there is no clear rational choice. Essentially, the problem falls outside of the domain of utility theory and alternative methods are needed, as discussed in Chapter 8.

4.7 Summary

In this chapter, a particular sub-class of imprecise probabilities was introduced. This method, called probability bounds analysis (PBA) and developed by other researchers, is slightly less expressive than general imprecise probability theory, but it is more intuitive and easier to compute with. The definition and interpretation of PBA and its model of uncertainty, called a p-box, were presented. Also, the relationship between p-boxes and intervals of expected utility, the creation of p-boxes, and the use of p-boxes in decision making were discussed. In the remaining chapters of the thesis, the value and using PBA in engineering design is explored.

CHAPTER 5:

COMPARING DIFFERENT METHODS FOR REPRESENTING UNCERTAINTY

In the preceding chapters, the nature of uncertainty was discussed, several models of uncertainty were considered, and one model selected as the most promising. This model was subsequently refined and explained in Chapter 4. The subject of the next three chapters of this dissertation is the question: *how useful is this model to engineering designers*? This is an important step in validating the answers advanced for motivating questions 1 and 2. If there is value in representing both imprecision and irreducible uncertainty using imprecise probabilities and p-box models, then this information combines with the theoretical arguments advanced in Chapter 2 through Chapter 5 to form a strong argument for the validity of the answers.

The first section of this chapter discusses the third motivating question of the dissertation: *how should engineers compare alternative models of uncertainty*? In Section 5.3, a specific experiment is described that compares two uncertainty models (best-fit, precise probabilities and p-boxes in PBA) in terms of practical value of the outcome of the design process. The remainder of the chapter, much of which was previously published in a conference paper (Aughenbaugh and Paredis 2005) and a forthcoming journal paper (Aughenbaugh and Paredis 2006a), is a discussion and explanation of the results of this experiment, as well as an analysis of the implications for engineering design. In Chapter 6 and Chapter 7, more general comparisons are made between PBA and a particular sensitivity analysis approach to accounting for imprecision.

5.1 Demonstrating the value of an uncertainty model

One way of showing the value of a model is by demonstrated that it is a generalization of another model, meaning that it can represent everything the other model can and then some. For example, it was already discussed that imprecise probabilities generalize precise probabilities in that precise probabilities are a specific case of imprecise probabilities. Another question is whether one model allows designers to rationally answer the same types of questions as another model, and in addition provides information that is not available using the other model. In Chapter 6, an argument is presented that probability bounds analysis generalizes specific aspects of sensitivity analysis procedure of decision analysis.

A second way to demonstrate the value of one model involves refocusing on the actual goal of engineering design: a profitable product. From this perspective, one model of uncertainty is better than another model is the first model enables the designer to create better designs than can be created using the second model.

Due to the obvious presence of uncertainty, it may be necessary to consider the worst case, best case, or average value of one model over another for a particular application. A procedure for making such comparisons using average value is presented in this chapter using the specific comparison of a best-fit, precise probability model to a p-box model in the context of limited statistical data using a design scenario described in the next section. The contributions of this chapter are thus both the general comparison method and the particular results of the comparison.

In the next chapter, a different approach is taken towards validating answers 1 and 2. This argument is built by showing how PBA generalizes a commonly used approach to uncertainty modeling in engineering design. This argument, unlike the one presented in this chapter, is therefore independent of any particular scenario. However, general applicability does not necessarily correspond to general value, hence the two chapters must be taken together in order to validate PBA and imprecise probabilities as an acceptable approach to uncertainty modeling in engineering design that is often preferable to other uncertainty models.

5.2 Example design scenario

Assume a decision maker (DM) needs to design a pressure vessel that is to contain 0.15 m^3 of gas under 7 MPa of pressure. Due to space limitations, certain maximum dimensions are imposed. The goal is to determine the dimensions (radius R, wall thickness t, and length L) of the vessel, shown in Figure 5.1, for which the overall utility, defined in Equation (5.1), is maximized. The vessel will be made of a new type of steel for which the yield strength is not well characterized. The material production process produces variations in the material properties such that the material yield strength is well modeled by a normally distributed random variable.

Because the material is new and testing is relatively expensive, variations in yield strength have only been measured in a set Σ of n independent tension tests, where n is a relatively small number due cost considerations. These tests can at best give an estimate of the true distribution, so in addition to inherent randomness (irreducible uncertainty), DM also faces imprecision—he or she cannot characterize the parameters of the random variable precisely. The number of samples n can be varied to explore different levels of imprecision, as discussed in Section 5.4.2.

Since the vessel will be used in a human-occupied location, the consequences of failure are significant. The DM accounts for the risk associated with vessel failure

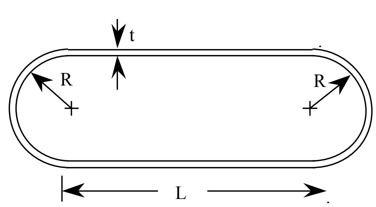


Figure 5.1. Pressure vessel schematic and design variables

explicitly in a utility function based on payoff in dollars and shown in Equation (5.1). This is a very simplified utility function, in which the single attribute of concern is net payoff, and the relationship between payoff and utility is assumed to be linear.

$$
\begin{aligned}
&U(\sigma_y, DV) = P_{selling} - C_{material} * volume(DV) - C_{failure} * \delta(\sigma_y, DV), \\
&\text{where:} \\
&\quad P_{selling} \equiv \text{selling price} = \$200 \\
&\quad C_{material} \equiv \text{material cost per volume} = \$8500/\text{m}^3 \\
&\quad \sigma_y \equiv \text{true yield strength of pressure vessel} \\
&\quad DV \equiv \text{design variables (radius, thickness, length)} \\
&\quad C_{failure} \equiv \text{cost incurred if vessel fails} = \$1,000,000 \\
&\quad \delta(\sigma_y, DV) \equiv \text{failure indicator} = \begin{cases} 0 \text{ if } \sigma_y \geq \sigma_{\max}(DV) \\ 1 \text{ otherwise} \end{cases}
\end{aligned}
\tag{5.1}
$$

The explicit inclusion of risk is seen more obviously if we consider the expected utility:

$$
\begin{aligned}
&EU(\sigma_y, DV) = P_{selling} - C_{material} * volume(DV) - E[C_{failure} * \delta(\sigma_y, DV)] \\
&where \\
&\quad E[C_{failure} * \delta(\sigma_y, DV)] = C_{failure} * P(\sigma_y < \sigma_{\max}(DV)) \\
&\qquad\qquad = \text{consequences * probability of failure} \equiv \text{ risk}
\end{aligned}
\tag{5.2}
$$

This formulation allows the DM to recognize that risk, and more specifically the pressure vessel's probability of failure (meaning the probability that the yield strength is less than the maximum stress in the walls of the pressure vessel), plays a very important part in design decisions. It thus seems important to characterize these probabilities appropriately.

5.3 Experiment comparing uncertainty models

The goal of the experiment is to compare the utility of the design solutions that result when different approaches for representing uncertainty are applied to the same design problem. The comparison is made possible in this experiment because it is assumed that overseeing the experiment is a supervisor who is in a state of precise information about the steel's material properties. From the supervisor's perspective, only irreducible uncertainty exists—uncertainty about the yield strength of the material that is precisely characterized by a normal distribution with a mean of 180 MPa and a standard deviation of 15 MPa. The supervisor can therefore determine precisely the dimensions of the pressure vessel that result in the maximum expected utility. This optimal design under the precise information is the benchmark for comparison of the other design approaches.

The general layout of the experiment is shown in Figure 5.2. The experiment consists of two DMs who are identical except in their method for modeling uncertainty: one uses approach A to model uncertainty about the yield strength (in this example a single best-fit normal distribution) and the other uses approach B to model uncertainty about the yield strength (in this example a p-box). The general experimental method is explained before including the details of specific experiment (shown in Figure 5.3) performed in this dissertation.

Both approaches start with the same information about the uncertain quantity (in this example the yield strength of the material). This information is a set Σ of n random samples (torsion test results) from the true distribution:

$$\Sigma = \{\sigma_{y_i}\}_{i=1}^{n} \tag{5.3}$$

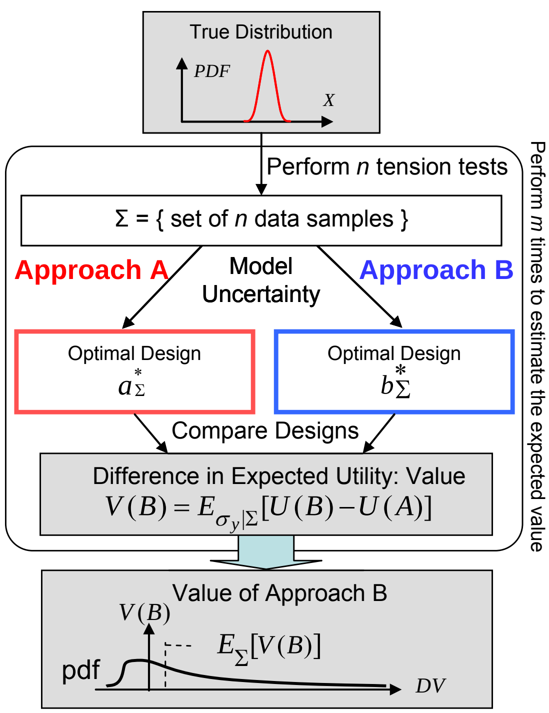


Figure 5.2. General experiment for comparing uncertainty models in engineering design decisions

(shaded actions are performed by a supervisor)

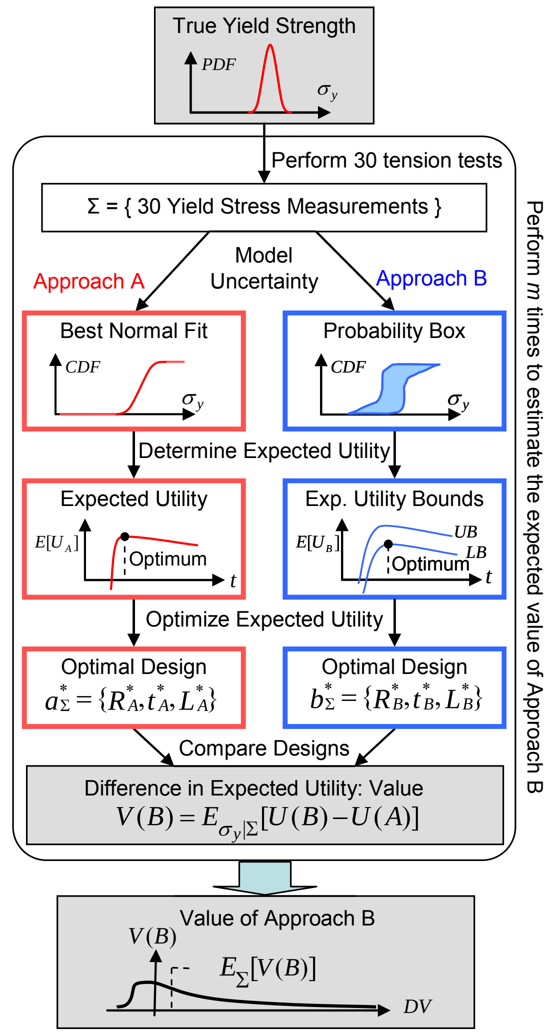


Figure 5.3. A computational experiment for determining the value of using imprecise probabilities

(shaded actions are performed by a supervisor)

Each DM models the uncertainty in these samples according to his or her own uncertainty model. They each then apply a decision policy that is appropriate for the assigned uncertainty model in order to select an optimal design, denoted as a_Σ^* for approach A and as b_Σ^* for approach B. The decision policies cannot be identical because the underlying models of the problem are different for each DM. For example, when imprecise probabilities are used, the expected utility of an alternative becomes an interval (see Section 4.6.1), so interval-based decision policies must be used. With the traditional best-fit approach, standard expected utility maximization can be used. Since the decision policies and the uncertainty models vary, one could think of the experiment as comparing design approaches rather than uncertainty models. The experiment is general in that it could even be repeated with both DMs using the same uncertainty model but different decision policies.

After the DMs have chosen their optimal designs, the supervisor compares each of the design solutions to determine the expected utility evaluated under precise information for each solution.

For approach A this is written as

$$\mathrm{E}_{\sigma_y|\Sigma}[U(\sigma_y, a_\Sigma^*)] \tag{5.4}$$

and for approach B as

$$\mathrm{E}_{\sigma_y|\Sigma}[U(\sigma_y, b_\Sigma^*)] \tag{5.5}$$

In order to compare the value of the two approaches, the supervisor, who has access to precise information, computes the difference in expected utility. Using the fact that the two approaches start with the same information (sample set Σ), the value of approach B over approach A can then be expressed as

$$\mathrm{V(B)} = \mathrm{E}_{\sigma_y|\Sigma}[U(\sigma_y, b_\Sigma^*) - U(\sigma_y, a_\Sigma^*)] \tag{5.6}$$

It is necessary to note that this value was for only one particular Σ—the set of yield strength measurements with which both designers start. Due to the randomness in Σ, one trial is not sufficient to judge the relative value of each uncertainty model; the supervisor needs to repeat the above experiment many times in order to determine which design approach performs best on average, over m different sample sets Σ. Mathematically, the expectation must be taken with respect to Σ in order to calculate the average expected value of approach B over A, written

$$\mathrm{E}_{\Sigma}[\mathrm{V(B)}] = \mathrm{E}_{\Sigma}[\mathrm{E}_{\sigma_y|\Sigma}[U(\sigma_y, b_{\Sigma}^{*}) - U(\sigma_y, a_{\Sigma}^{*})]] . \tag{5.7}$$

The addition of the word average emphasizes that this quantity is the expectation over the samples of the expected utility of particular design solutions.

In this section, a method for comparing design approaches was presented. A final caveat on the method is that it can only compare uncertainty models that accept the same type of input data (specifically, observed statistical data). For example, it could not be used to compare a method that requires pure expert opinion with one that requires statistical data samples. The remained of this section presented the customized experiment that compares a best-fit precise probability approach (approach A) with a p-box based approach (approach B).

5.3.1 Design using approach A: precise normal fit

Designer A does not have access to precise information, but instead only has access to the set Σ of n data samples. Because the designer does not know the true distribution of σ_y, he or she must make an approximation, denoted $\tilde{\sigma}_y(A,\Sigma)$. In this approximation, the representation of $\tilde{\sigma}_y$ depends on both the approach, in this case A, and the observed random sample Σ. Designer A represents the uncertainty as a normal distribution, using the sample mean and sample variance as unbiased estimates of the true mean and true variance, respectively. This yields the probabilistic model

$$\tilde{\sigma}_y(A,\Sigma) \sim N(\tfrac{1}{n}\sum_{i=1}^{n}\sigma_{y_i}, \tfrac{1}{n-1}\sum_{i=1}^{n}(\sigma_{y_i} - \tfrac{1}{n}\sum_{i=1}^{n}\sigma_{y_i})^2). \tag{5.8}$$

Designer A therefore chooses design variables $a_\Sigma = \{R_A, t_A, L_A\}$ that maximize the estimated expected utility given his or her information about the randomness. The expected utility is only estimated because designer A does not have access to a precise characterization of the random variable σ_y. The expected utility maximization results in the optimal design using approach A given samples Σ, denoted:

$$a_\Sigma^* = \arg\max_{a_\Sigma}(\mathrm{E}_{\tilde{\sigma}_y|\Sigma}[U(\tilde{\sigma}_y(A,\Sigma), a_\Sigma)]) \tag{5.9}$$

5.3.2 Design using approach B: imprecise probabilities

Designer B takes a different approach for capturing the uncertainty in the yield strength. Specifically, designer B represents the uncertainty in σ_y by $\tilde{\sigma}_y(B,\Sigma)$, where the uncertainty in $\tilde{\sigma}_y(B,\Sigma)$ is modeled using a parameterized p-box rather than a precise normal distribution, such that $\tilde{\sigma}_y(B,\Sigma) = \boxed{\sigma_y}^{p}$. The p-box is constructed using 95% confidence intervals on the mean and variance of the yield strength (see Section 4.3), such that with $\alpha = 0.05$:

$$[\underline{\mu}, \overline{\mu}] = [\hat{\mu} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}},\ \hat{\mu} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}] \tag{5.10}$$

$$\left[\underline{\mathrm{var}}, \overline{\mathrm{var}}\right] = \left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right]. \tag{5.11}$$

The estimated expected utility under design approach B is defined as

$$\mathrm{E}_{\tilde{\sigma}_y|\Sigma}[U(\tilde{\sigma}_y(B,\Sigma), b_\Sigma)] \tag{5.12}$$

where $b_\Sigma = \{R_B, t_B, L_B\}$ is the designer's chosen design action given sample set Σ. Because the p-box expresses a range of possible distributions, the expected utility is no longer a crisp number but rather an interval defined by lower-bound $\underline{E}$ and upper-bound $\overline{E}$ (as described in Section 4.6.1), such that

$$\mathrm{E}_{\tilde{\sigma}_y|\Sigma}[U(\tilde{\sigma}_y(B,\Sigma),b_\Sigma)] = [\underline{E},\overline{E}]. \tag{5.13}$$

Because the expected utility is now an interval, the designer cannot choose design variables b_Σ that maximize the expected utility in the traditional sense, as mentioned in Section 4.6 and discussed in more detail in Chapter 8. In this experiment, a conservative best worst-case, or Γ-maximin policy (Berger 1985) is used. Designer B therefore chooses the design action b_Σ that has the highest lower bound $\underline{E}$ on the expected utility. This results in an optimal design decision using approach B given the observed samples Σ, denoted:

$$b_\Sigma^* = \arg\max_{b_\Sigma}(\underline{E}) = \arg\max_{b_\Sigma}(\underline{\mathrm{E}}_{\tilde{\sigma}_y|\Sigma}[U(\tilde{\sigma}_y(B,\Sigma),b_\Sigma)]). \tag{5.14}$$

5.3.3 Supervisor's design under precise information

In addition to approach A and approach B, the experiment's supervisor can create a design using the true distribution, since he or she is in a state of precise information. The supervisor therefore knows precisely that $\sigma_y \sim N(180MPa,(15MPa)^2)$. If approach is defined as approach K (not shown in Figure 5.3), the supervisor then chooses the design variables $DV = k = \{R_K, t_K, L_K\}$ such that the expected utility $\mathrm{E}_{\sigma_y}[U(\sigma_y,k)]$ is maximized. This leads to the optimal design under precise information:

$$k^* = \arg\max_{\mathrm{k}}(\mathrm{E}_{\sigma_y}[U(\sigma_y,k)]). \tag{5.15}$$

This optimal design, with expected utility denoted $\mathrm{E}[U(k^*)]$ for brevity, serves as the baseline for comparison because no other approach can yield an average higher expected utility across many repetitions m.

5.4 Experimental results

The computational experiment was repeated for many different initial sample set sizes, that is many different n. For each level of n, the design process was repeated for $m = 100,000$ different initial sample sets Σ in order to determine the average performance of the two methods. The results for the particular sample size $n = 25$ are presented first, and then the results over varying values of n, which represent different levels of imprecision, are discussed.

5.4.1 Value of using imprecise probabilities for 25 samples of the true yield strength

The experiment is first conducted with a sample set Σ of size $n = 25$, meaning the designers are given the results of 25 independent yield stress tests. As measured by the supervisor using precise information and averaged over $m = 100,000$ initial sets, approach B on average yields designs with greater expected utility than approach A. Specifically, using standard statistical analysis, the 95% confidence interval (CI) on the value of approach B over approach A shown in Equation (5.16) is found.

$$\text{95\% CI on } V(B) \text{ is [\$22, \$26]} \tag{5.16}$$

To put this result in perspective, the expected utility of the supervisor's design, which is the best possible because it is designed under a state of precise information, is $\mathrm{E}[U(k^*)] = \$104$. Thus the CI on the expected value of approach B over A can also be expressed as [21%, 25%] of the optimal utility $\mathrm{E}[U(k^*)]$. This is a substantial deviation that suggests that there is value in using the p-box approach for this design problem. The average expected utilities realized under approach A and B are given in Equations (5.17) and (5.18) .

$$\text{Approach A: } \mathrm{E}_\Sigma[\mathrm{E}_{\sigma_y|\Sigma}[U(\sigma_y, a_\Sigma^*)]] = \$34 \tag{5.17}$$

$$\text{Approach B: } \mathrm{E}_\Sigma[\mathrm{E}_{\sigma_y|\Sigma}[U(\sigma_y, b_\Sigma^*)]] = \$58 \tag{5.18}$$

The total deviations from optimal ($70 for approach A and $46 for approach B), coupled with the relative value of approach B over A, indicate that B is a better approach at this level of imprecision. In the next section, the variation of these results at different levels of imprecision is examined

5.4.2 Variation of value with level of imprecision

The previous discussion dealt with a fixed sample set size of $n = 25$ material strength tests. While those results demonstrated that it was valuable to use the p-box approach in that case, a more general result is desirable. By varying the number of material strength tests n, one can vary the imprecision of the characterization. The supervisor's design yields the best possible expected utility $\mathrm{E}[U(k^*)]$, and hence the designs of designers A and B can at best equal it. In Figure 5.4, the percent deviation from this best-possible expected utility for approach A and approach B for different values of n are shown in log-log scale. Because this is a deviation from best-possible, the smallest absolute value is desirable. Hence, examining Figure 5.4, smaller is better.

When the imprecision is large, approach B performs significantly better than approach A. For example, at a sample size of 10, a 95% CI on the value of approach B over A is [520%, 570%] of $\mathrm{E}[U(k^*)]$. There is no doubt that this value is significant.

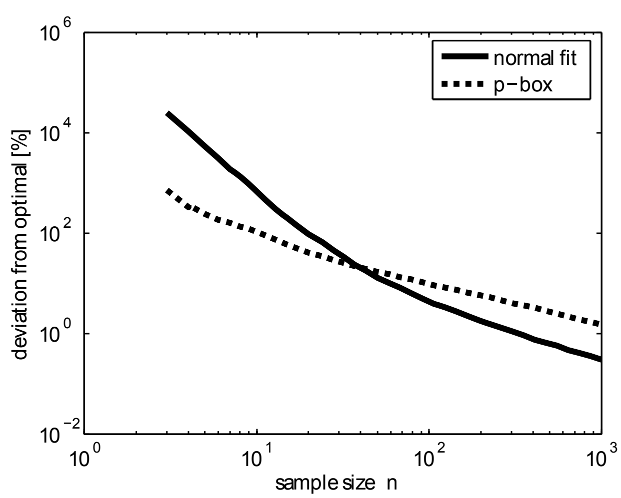


Figure 5.4. Variation of value with imprecision

Around sample size 40, the two approaches yield similar results. Past this point, approach A performs better, but not by much. For example, for a sample size of 100, the 95% CI on the value of approach B is [-5.5%, -5.3%] of $\mathrm{E}[U(k^*)]$.

In summary, as the imprecision increases, the value of approach B over approach A increases significantly. As the imprecision approaches zero, approach A becomes only slightly better than B. For different design problems, the designer will not necessarily know where the two curves cross. Thus, unless the designer is sure *a priori* that the consequences of the imprecision are insignificant, the results of this computational experiment suggest that it is valuable to explicitly represent the imprecision in the available characterization of uncertainties by using imprecise probabilities.

5.4.3 Explanation of results

In this section, additional insight into the results of the previous two sections is presented. The results for 100,000 different sample sets Σ of sizes n=10, 25, and 100 are shown in histograms in Figure 5.5. The x-axis of the histogram in Figure 5.5 is the expected value of approach B over A, denoted as $\mathrm{E}_{\sigma_y|\Sigma}[U(\sigma_y, b_\Sigma^*) - U(\sigma_y, a_\Sigma^*)]$.

In many cases, approach B yields a design with a lower utility than approach A due to the extra material costs of the more conservative design. However, in some cases, approach B yields a design with a much higher expected utility. The tails of the distribution for $n = 10$ and $n = 25$ extend much farther to the right than shown in the figure. For example, the maximum expected value of approach B over A seen in any trial of $n = 10$ was $220,000. Results such as these skew the overall distribution such that on average, approach B yields a design with a higher expected utility than approach A yields. As the imprecision decreases, both the skewness of the distribution and the value of approach B over A decrease.

The results can also be understood in terms of the expected utility curves used in the experiment. In Figure 5.6 and Figure 5.7, the expected utility is illustrated as a function of wall thickness t for two different samples, Σ_1 and Σ_2 respectively, both of size $n = 25$. The four curves shown in each figure are:

1. Estimated expected utility under approach A: $\mathrm{E}_{\tilde{\sigma}_y|\Sigma}[U(\tilde{\sigma}_y(A,\Sigma),t)]$

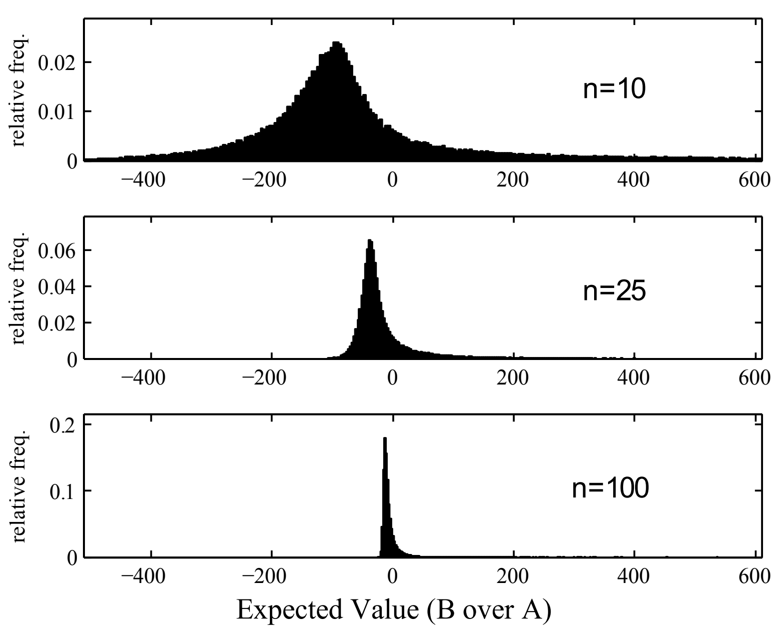


Figure 5.5. Histogram of value of p-box approach

2. Estimated lower bound on expected utility under approach B: $\underline{\mathrm{E}}_{\tilde{\sigma}_y|\Sigma}[U(\tilde{\sigma}_y(B,\Sigma),t)]$
3. Estimated upper bound on expected utility under approach B: $\overline{\mathrm{E}}_{\tilde{\sigma}_y|\Sigma}[U(\tilde{\sigma}_y(B,\Sigma),t)]$
4. The true expected utility in a state of precise information: $\mathrm{E}_{\sigma_y}[U(\sigma_y,t)]$

Designers A and B choose a thickness (t_A^* and t_B^* respectively) that is optimal according to their estimated expected utilities. Their estimates are in general not equivalent to the true expected utilities. Therefore, the expected utility actually realized by a particular design is not reflected in their estimates, but rather in the true curve $\mathrm{E}_{\sigma_y}[U(\sigma_y,t)]$, which is known only to the supervisor.

In Figure 5.6 (based on sample set Σ_1), the true curve is between the curve from approach A and the upper bound from approach B. By noting t_A^* and t_B^*, one can read off the true expected utility evaluated under precise information for each approach from the truth, $\mathrm{E}_{\sigma_y}[U(\sigma_y,t)]$. For this particular sample set Σ_1, the expected utility realized from approach A is about \$20 higher than the expected utility realized from approach B. This means for sample Σ_1 the relative value of approach B is negative: V(B) = -\$20, indicating approach A performs better.

In Figure 5.7 (based on sample set Σ_2), the true curve is near the lower bound of approach B. For this particular sample set Σ_2, the expected utility from approach B is about \$70 more than that from approach A, so the relative value of approach B for sample Σ_2 is V(B)=\$70.

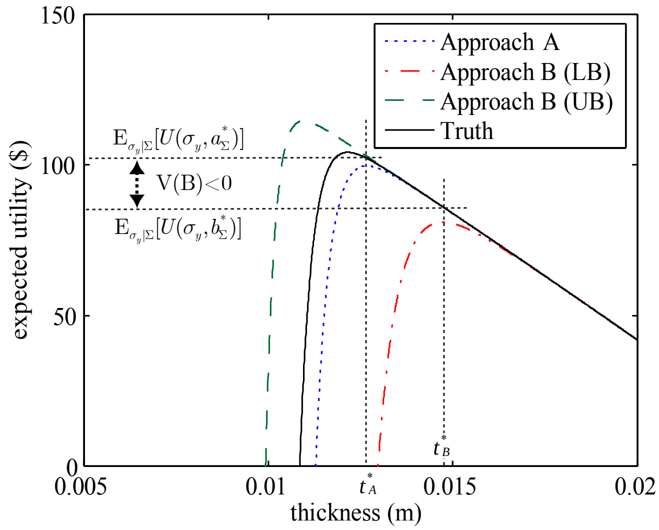


Figure 5.6. Example expected utility functions, V(B)<0

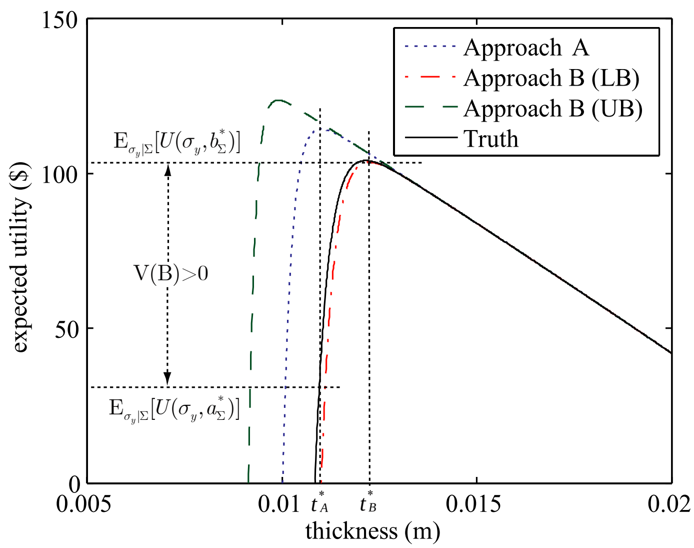


Figure 5.7. Example expected utility functions, V(B)>0

The preceding results are for two representative cases. The overall results of the experiment, discussed previously, indicate that, on average, the latter case dominates. Approach A is more likely to overestimate the true material strength, and when it does, the consequences are disastrous — a high probability of failure. Approach B is more conservative, resulting in higher material costs, but, on average, these material costs are offset by the reduced failure costs. As the sample size increases, the best-fit normal

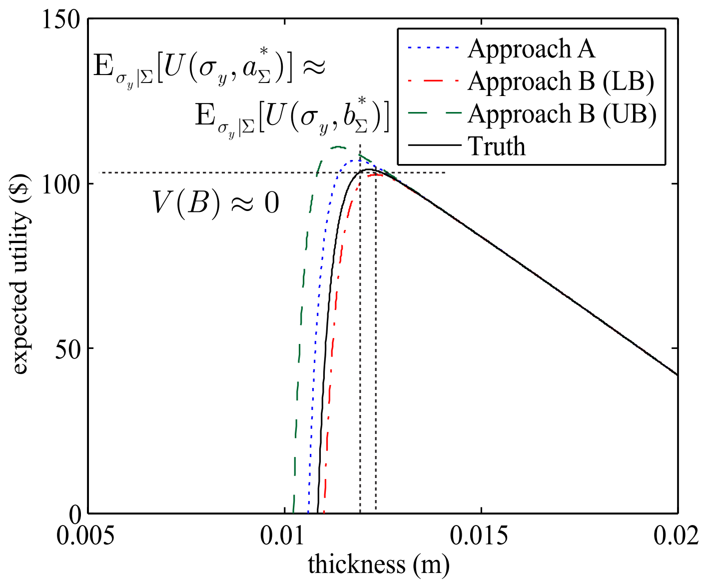


Figure 5.8. Example expected utility functions, V(B)=0

distribution of approach A becomes, on average, closer to the true distribution, such as the example sample Σ_3 shown in Figure 5.8. For this sample, the optimal designs of the three approaches converge, and therefore yield similar utilities. Thus one can see that when the imprecision is small, the value of approach B over A is near zero.

5.4.4 Summary of results

For this design problem, the experimental results indicate that when the imprecision is large, approach B (using imprecise probabilities) performs significantly better than approach A (using precise probabilities). When the imprecision is small, the difference between the two approaches is insignificant. This computational experiment has therefore demonstrated that there are scenarios in which it is valuable to explicitly represent the imprecision in the available characterization of uncertainties by using imprecise probabilities.

5.5 Discussion of experimental results

So far in this chapter, a specific and simplified design problem and computational experiment that demonstrates the value of using imprecise probabilities in engineering design has been presented. There are other issues to consider, such as computation cost and decision policies. Additionally, it is always useful to explore the results in more general problems.

5.5.1 Computational costs

In this paper, the potential value of a design method that uses imprecise probabilities has been demonstrated. What has heretofore been referred to as value is really the gross value or benefit. According to the principles of information economics (Aughenbaugh, et al. 2005), designers should really care about the gain, or net value, which is defined as the difference between gross value (benefit) and cost, including computational cost.

In this example, the computational costs for both approaches were small due to the structure of the problem. In reality, even the best-fit probabilistic approach often will require Monte Carlo analysis or related methods in order to calculate expected values (Robert and Casella 1999). These methods have an inherently high computational cost, even when used in combination with surrogate modeling (Myers and Montgomery 1995, Simpson, et al. 2001) to reduce the computational burden. An obvious way to compute with p-boxes is to perform a second-order Monte Carlo simulation, in which an outer loop is added around the traditional loop (Vose 2000). Such methods can be extremely expensive in terms of computations.

However, the DBC method discussed in Section 4.5.1 can be used instead. Ferson and Ginzburg (Ferson and Ginzburg 1996) show that these methods are on average much less computationally expensive than second order Monte Carlo methods. However, additional work is needed to compare the actual additional computational burden of PBA with the additional benefits from using PBA in complex problems. Because the nature of the computations is so different, straightforward comparisons are often difficult.

An added benefit of PBA is the flexibility of the DBC algorithm. In Monte Carlo analysis, engineers often assume specific dependencies between variables as a matter of convenience that enables the use of standard statistics. This may understate the true imprecision in the available characterizations of uncertainty and may have serious consequences. PBA provides methods (Ferson, et al. 2002b) that can propagate uncertainty under various conditions of dependence and correlation, including the extremes of fully known and completely unknown dependencies.

Despite this promise, the application of these techniques has yet to be explored in complex engineering design problems. It is therefore unclear how well PBA can propagate uncertainty from many sources, or how well PBA can be integrated with more complex design tools and models, such as discrete-event simulations and optimization methods. These issues need to be addressed before PBA can be put to use in general engineering design problems.

5.5.2 Decision policies and preferences

In this experiment, the relatively high consequences of vessel failure, in comparison to material costs, makes failure avoidance a key driver in the design. The Γ-maximin decision policy used in this experiment is conservative compared to a normal fit approach. Because it uses the lower bound on the expected utility, it is actually the most conservative policy that is consistent with the available evidence.

Other decision policies (such those that account for shared uncertainty such as maximality (Walley 1991) and E-admissibility (Levi 1974), discussed in Chapter 8, or the arbitrary choice policy of the Hurwicz criterion (Arrow and Hurwicz 1972)) are possible and may perform better in different circumstances. For example, if instead of the lower bound on expected utility one uses the midpoint between the upper and lower bounds (a Hurwicz policy with criterion of 0.5), the results change. For example, the two curves in

Figure 5.9 cross at 45 samples using the Hurwicz policy, as compared to 40 in Figure 5.4 with the Γ-maximin policy.

Part of the motivation of using a policy such as the Hurwicz criterion is that it allows the designer's preferences (in the form of the utility function) to influence the resolution of imprecision. A best-fit approach ignores the nature of the DM's preferences. The distinction between the two approaches is illustrated with the following example, as shown in Figure 5.10. When the utility function is skewed due to a high cost of consequences in the risk term, such as in this example, the maximum of the midpoint of the expected utility bounds is much closer to the maximum of the lower bound than to the maximum of the best-fit solution, as shown in. By explicitly accounting for imprecision, a more appropriate and quantitative performance/risk tradeoff can be performed.

An obvious question is: *why is the midpoint solution not similar to the best-fit solution*? The answer highlights an important advantage of a p-box approach over a best-fit approach. The midpoint is applied to the expected utility distributions, whereas the best-fit is done on the parameters of the uncertainty model. In general, the best-fit approach leaves the decision-maker to deal with imprecision qualitatively and outside the

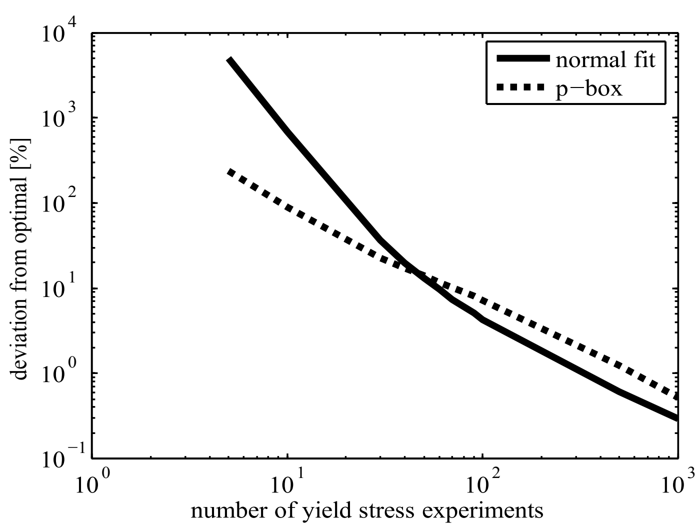


Figure 5.9. Variation in value with imprecision for midpoint policy

uncertainty formalism, while the p-box approach quantifies the imprecision explicitly in utility curves. In a scenario in which only one decision is to be made regarding a particular uncertain quantity and imprecision is high, it is clearly useful to model uncertainty using p-boxes. However, identifying this situation is more complicated.

Arbitrary policies are limited in their applicability in engineering design because it is difficult to apply them (other than Γ-maximin) in a way that is rationally consistent across multiple decisions that involve the same uncertain quantities. This is not necessarily surprising since the Hurwicz policy essentially breaks a tie between two alternatives arbitrarily. However, it does suggest a need for future work in decision making under uncertainty. Fortunately, other policies are available for rationally and consistently reducing the set of alternatives even when the alternatives are imprecisely characterized. These policies are the subject of Chapter 8. The question remains as to how to make a final decision when the intervals of alternatives overlap and there is no economic way to further reduce the imprecision.

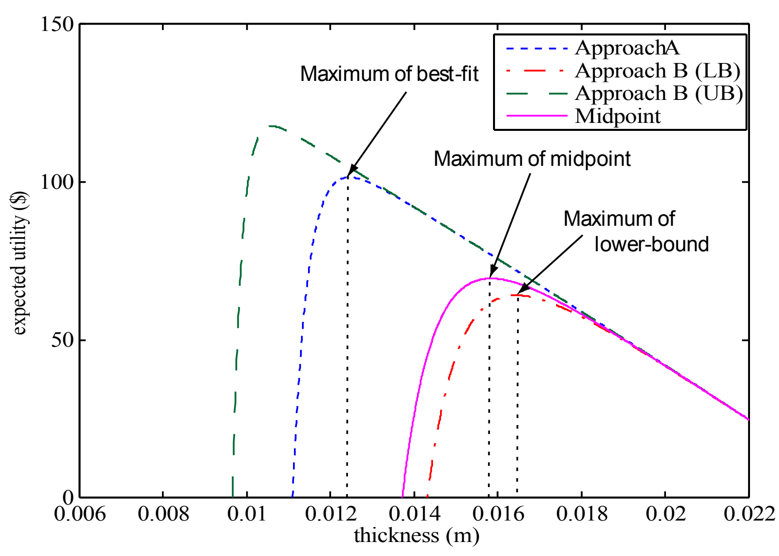


Figure 5.10. Midpoint policy results

5.6 Summary

In this chapter, an experiment was developed that allows for two uncertainty models to be compared in terms of the quality of the final product that results from the design process. This method is part of the answer to motivating question 3 of the dissertation, which asks how should engineers compare alternative models of uncertainty? This method is then used to help validate the answers advanced in response to motivation questions 1 and 2.

Specifically, the pressure vessel design example and experiment in this chapter demonstrates that when the designers only have access to a small set of sample data, a design approach that uses imprecise probabilities to model uncertainty in design decisions leads on average to better designs than a purely probabilistic approach that requires precise probabilities.

It can then be concluded that in some design problems, it is valuable to represent the imprecision in the available characterization of uncertainties explicitly by using imprecise probabilities. In the next chapter, properties of PBA that are more general in their applicability and value are described, thus moving past the context of a specific, simple design example.

CHAPTER 6:

PBA AS A GENERAL APPROACH TO SENSITIVITY ANALYSIS

In the previous chapters, several discussions and results regarding imprecision in engineering design have been presented. In Chapter 2 and Chapter 3, the general motivation for considering imprecision explicitly was developed. In Chapter 4, the probability bounds analysis (PBA) approach was introduced. The potential value of using PBA was demonstrated in Chapter 5 in the context of a high-risk, compromise decision.

In this chapter, acceptance of the existence of imprecision is assumed. People often recognize their lack of knowledge implicitly, even when they use formalisms such as precise probabilities that do not recognize this imprecision. They often perform "what-if" analyses that compare the assumed scenario to different, but often closely related scenarios. For example, it is common in decision analysis to perform a sensitivity analysis as part of the decision process (Clemen 1996). One purpose of a sensitivity analysis is to examine whether the decision is sensitive to lack of knowledge. This is achieved by exploring whether the optimal decision action changes as the uncertain parameters are varied across some neighborhood of the decision-maker's best-guess, or base values. Bayesian decision theorists often perform Bayesian sensitivity analysis in which various prior distributions are examined (Berger 1985). By performing a sensitivity analysis, the decision-maker is in essence checking whether his or her lack of knowledge affects what he or she decides.

Due to computational costs and other factors, decision-makers are normally limited to performing one-way (and occasionally two-way) sensitivity analyses. In a one-way analysis, each uncertain parameter is varied over a range with the others fixed at their base values (Clemen 1996). This analysis neglects interactions between the uncertain parameters. In order to be absolutely sure that a decision is not sensitive to imprecision, a decision-maker would need to perform an all-way analysis, which means comparing all combinations of reasonable values across all uncertain parameters. In this chapter, it is shown that an uncertainty analysis using PBA generalizes the process of an all-way effects analysis for selection decisions in a computationally efficient manner.

In Sections 6.1 and 6.2, existing sensitivity analysis approaches are discussed. In Section 6.3, it is explained how PBA generalizes sensitivity analysis; it provides measures of the sensitivity of a decision to imprecision, measures that are more conservative and more rigorous than those provided by traditional methods. In Section 6.4, some limitations of PBA-based sensitivity are presented, and an extension of PBA-based sensitivity analysis is proposed in Section 6.4.3. Finally, the use of PBA as a sensitivity analysis method is demonstrated and discussed using the example of the design of an environmentally benign oil filter in Chapter 7.

6.1 Traditional sensitivity analysis approaches

The notion of "sensitivity analysis has been used to describe various procedures and specific goals in engineering design. As such, there is no universal definition of what a sensitivity analysis is or should be. At the highest level of abstraction, a sensitivity analysis is the study of how certain things influence other things. More specific to the context of engineering design, a sensitivity analysis is the quantitative study of how the inputs to a model affect the outputs (Ferson, et al. 2004a), where a model is defined broadly and includes all functions, calculations, and simulations. Ferson and coauthors (Ferson, et al. 2004a) also note that there are fundamentally two reasons for conducting a sensitivity analysis: to understand the reliability of conclusions and inferences drawn from an analysis (which will be called sensitivity analysis for *decision robustness* in this dissertation) and to focus future information collection efficiently on those aspects to which the problem is most sensitive (which will be called sensitivity analysis for *information prioritization* in this dissertation).

A lot of research in engineering design has focused on probabilistic sensitivity analysis (Homma and Saltelli 1996, Sobol 2001, Chen, et al. 2004, Felli and Hazen 2004, Liu and Chen 2006). These methods focus on identifying the largest contributors to probabilistic effects in the output. For example, one type of approach attempts to determine how much each input contributes to the statistical variance of the output. These methods are not exploring sensitivity to a lack of information, but rather sensitivity to a physical reality of the problem or process being modeling. This analysis can provide important insight during the design process, such as indicating that a manufacturing process needs to be improved so that the variance of the thickness of parts is reduced to an acceptable level. However, such analyses do not help a DM explore the robustness of his or her decisions to the available information, which is the focus of this dissertation.

Another form of sensitivity analysis looks at the partial derivatives of a function (Ferson and Tucker 2006 (in preparation)). For example, if a model is represented by Equation (6.1), it is possible to express the sensitivity of the result to each input parameter in terms of partial derivatives, as shown in Equations (6.2)-(6.4):

$$f(a,b,c) = ab + a^2 c \tag{6.1}$$

$$\frac{\partial f}{\partial a} = b + 2ac \tag{6.2}$$

$$\frac{\partial f}{\partial b} = a \tag{6.3}$$

$$\frac{\partial f}{\partial c} = a^2 \tag{6.4}$$

The partial derivatives represent the local sensitivity of the function to the three independent variables. The sensitivity analysis is local in that it must be evaluated at some point $\{a,b,c\}$, and the sensitivities, like all partial derivatives, are only locally valid approximations. The comparison of the partial derivatives is complicated because the units are not consistent, since in general a, b, and c have different units. Nevertheless, the partial derivatives do give some insight into the impact that different input parameters have on the output.

Sensitivity measures made using partial derivatives capture an inherent quality of the problem, but they do not take into account the available information. For example, if the DM knows a, b, and c perfectly, then these derivatives are irrelevant to information management, because there would be no deviation from the perfectly known nominal values. A more useful analysis would explore the effects of imprecision on the output. For example, if the available knowledge is only that $a=[6,17]$, $b=2$, and $c=[1,2]$, then it is pretty clear from Equation (6.1) that the output is very sensitive to the imprecision in a, but the partial derivatives make it appear as if the greatest sensitivity is to c. It appears that a more complete approach is needed.

Leamer (1990) has defined a global sensitivity analysis as a systematic study in which "a neighborhood of alternative assumptions is selected and the corresponding interval of inferences is identified." As Ferson and Tucker note (2006 (in preparation)), such studies can be done in at least two obvious ways. In the first approach, the neighborhoods of consideration are expressed as bounds, thereby forming intervals of imprecise parameters. In the second, probabilities are assigned across the neighborhood, thus making the parameters probabilistic.

As established in the preceding chapters, the perspective adopted in this dissertation is that imprecision is expressed well using intervals, and thus the first approach is taken in this chapter. In fact, the use of probabilities to describe imprecision is in this instance contradictory to the decision robustness goal of sensitivity analysis, which is to identify the reliability of the inferences (or outputs) drawn from the assumed inputs. Adding an additional assumption of information about probabilities over the imprecision adds information, rather than explores the sensitivity of the true lack of information. The use of intervals is consistent with traditional decision analysis approaches to sensitivity analysis, as discussed in the following section.

This section is closed with a brief disclaimer. Since there is no universal agreement on what a sensitivity analysis is or how it should be performed, there is no universal standard for comparison. In the following section, a specific type of sensitivity analysis is described. The subsequent sections then discuss how the propagation and analysis of uncertainty using PBA generalizes this type of sensitivity analysis, providing information and rigor that is not available with the traditional approach.

6.2 Sensitivity analysis in decision analysis

Decision analysis is a discipline that studies procedures, tools, and frameworks for transforming problems that are difficult to understand, solve, or explain into problems that are more readily understood and solved (Howard 1988a, Clemen 1996). To distinguish the approach discussed in this section from alternative approaches, it will be referred to as decision analysis with sensitivity analysis, abbreviated as DASA in the remainder of this dissertation.

Clemen (1996, page 6) describes the DASA process with the following steps:

1. Identify the situation and understand objectives
2. Identify alternatives
3. Decompose and model the problem structure, uncertainties, and preferences
4. Choose the best alternative
5. Perform sensitivity analysis
6. Decide if further analysis is needed

7. Implement the chosen alternative

Steps 1-3 result in the specification of the decision model described in Section 3.6. Traditional decision analysis does not explicitly recognize imprecision when choosing the best alternative in Step 4, so the analysis is performed using best-guess, base values for all of the imprecise parameters.

6.2.1 Choosing the best alternative

Recall that the imprecise parameters are a subset of the uncertain parameters. Any probabilistic uncertainties are retained. Returning to the example of Section 3.6, if ${}_2x \sim N({}_1x, \sigma^2)$, then ${}_2x$ will still be treated probabilistically, but any imprecision in ${}_1x$ must be reduced to a base value, for example ${}_1\hat{x}$. More generally, the state vector $s = \{{}_1x, \ldots, {}_kx, {}_{k+1}x, \ldots, {}_2x\}$ is replaced with $\hat{s} = \{{}_1\hat{x}, \ldots, {}_k\hat{x}, {}_{k+1}x, \ldots, {}_nx\}$, such that ${}_1\hat{x}, \ldots, {}_k\hat{x}$ are all scalar values.

With ${}_1\hat{x}, \ldots, {}_k\hat{x}$ fixed to base values, an expected utility maximizing action a^* can be chosen using the precisely defined probabilities over the probabilistic parameters ${}_{k+1}x, \ldots, {}_nx$. Step 5 then considers the sensitivity of the optimality of a^* to the actual imprecision in the available information.

6.2.2 Performing a basic one-way sensitivity analysis

DASA assumes that the DM can specify a feasible range for the imprecise parameters, such as given by Equation (6.5).

$$ {}_ix = [{}_i\underline{x}, {}_i\overline{x}] \text{ for i=1...k} \tag{6.5} $$

The sensitivity analysis step of the decision process explores how moving the values of the ${}_ix$'s across their ranges affects the optimal decision. One convenient way of performing a sensitivity analysis for a selection decision is to evaluate the sensitivity of the decision outcome graphically using tornado diagrams (Howard 1988a, Eschenbach

1992, Clemen 1996). A tornado diagram allows a decision maker to perform a one-way sensitivity analysis--that is, to explore the effects of uncertain parameters one at a time.

A simple tornado diagram, such as shown in Figure 6.1, compares a single action (for example a_1) to another action (for example a_2), where it is assumed that action a_2 is entirely robust to the imprecision; that is, the expected utility of a_2 is completely independent of ${}_1x,\ldots,{}_kx$. The perfectly known expected utility of action a_2 is represented by the large vertical, dotted line on the tornado plot. Also displayed on the tornado diagram is a bar for each imprecise parameter, where each bar represents the range of the expected utility of action a_1 across the range of an imprecise parameter. The bars are created by taking one imprecise parameter ${}_ix$ and varying it from its lower limit ${}_i\underline{x}$ to its upper limit ${}_i\overline{x}$ with all other parameters held constant at their base values. This is repeated for each parameter. The bars are then reordered from largest (at the top) to smallest (at the bottom), thus creating the shape that gave rise to the name “tornado diagram."

If any of the bars of the tornado plot intersect the dotted line corresponding to the expected utility of a_2, then the decision is sensitive to the imprecision. Recall that each bar represents the range of expected utilities that can be realized for action a_1 depending on where in the imprecise region ${}_ix = [{}_i\underline{x}, {}_i\overline{x}]$ the true value of ${}_ix$ lies. When a bar

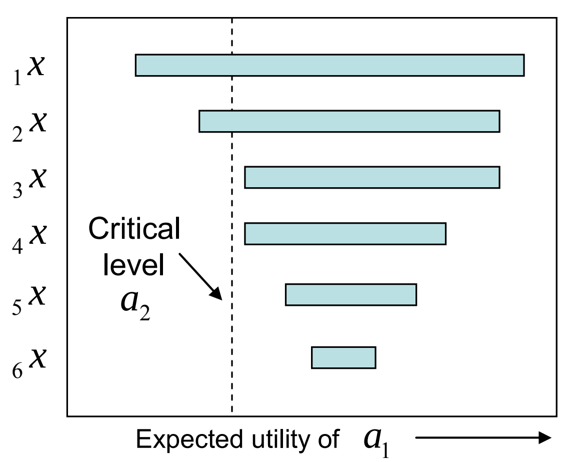


Figure 6.1. Sample tornado diagram, one imprecisely known alternative

crosses the dotted line, it means that depending on where the true ${}_i x$ lies, either a_1 or a_2 can yield a higher expected utility.

6.2.3 Performing a more advanced one-way sensitivity analysis

Tornado diagrams can be generalized to situations with more than two alternative actions, including cases in which there are multiple alternatives whose expected utilities are functions of imprecise parameters, although this is not addressed explicitly in the existing literature. Consider first the case in which there are two actions whose expected utilities depend on imprecise parameters. There are two ways to consider this problem. The first is to draw bars on a tornado plot for each alternative and each imprecise parameter. If the bars for the two alternatives overlap for the same parameter, then the decision is one-way sensitive to that imprecise parameter; overlapping intervals imply that the decision depends on the resolution of the imprecision, as discussed in Section 4.6.3.

There are two limitations to this straightforward method. First, such comparisons become graphically confusing for more than two alternatives. Second, this method does not account for shared uncertainty.

Consider the imprecise parameter ${}_i x = [{}_i \underline{x}, {}_i \overline{x}]$, and assume that ${}_i x$ represents the ambient temperature in which a system will operate. Obviously as ${}_i x$ is varied within its range, the expected utility of each action (a_1 and a_2) varies. However, the movement of the expected utilities is in a sense coordinated. Assuming that the action chosen has no effect on the resolution of the imprecision in ${}_i x$, then both actions would face the same true value of ${}_i x$—i.e., the same ambient temperature. This is a generalization of act-state independence, a probabilistic concept that states that $P({}_i x \mid a_i) = P({}_i x)$, meaning the probability of future states of the world are independent of the DM's actions. The expected utilities of two actions should therefore be compared using the same values for the ${}_i x$ under consideration in the sensitivity analysis.

For example, the expected utility for a_1 assuming ${}_i x = {}_i \underline{x}$ should be compared to the expected utility of a_2 also assuming ${}_i x = {}_i \underline{x}$. A comparison of a_1 assuming ${}_i x = {}_i \underline{x}$ to a_2 also assuming ${}_i x = {}_i \overline{x}$ provides no information, because in general ${}_i x = {}_i \underline{x}$ and ${}_i x = {}_i \overline{x}$ are contradictory. It thus only makes sense to construct a tornado plot of the differences of expected utility ($E[u(a_1) - u(a_2) \mid {}_i x]$) while considering this shared imprecision, as shown in Figure 6.2. In this approach and as portrayed in the figure, the comparison is made to a zero difference in expected utility. If a bar of the diagram crosses zero, it means that with the given imprecision, it is not clear whether alternative a_1 or a_2 yields the higher expected utility, and hence the decision is one-way sensitive to the parameters for which the bars cross zero. The distinction "one-way sensitive" is made to emphasize that the sensitivity was identified using a one-way, or one-factor-at-a-time analysis.

If a decision is found to be sensitive to imprecision, the DM has a few choices. One of these is to continue collecting information to reduce the imprecision, a topic addressed in more detail in Chapter 8. The DM can also conclude that the sensitivity is small enough that he or she does not care enough to spend resources to further reduce the sensitivity. The emphasis in this chapter is merely on the identification of the sensitivity.

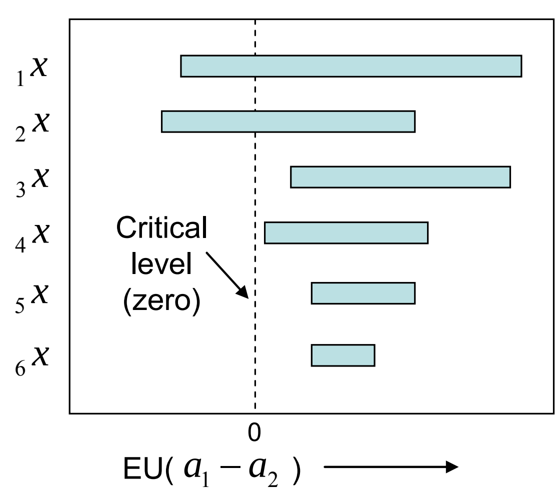


Figure 6.2. Sample tornado diagram, comparing alternatives

6.2.4 Beyond one-way sensitivity analysis

A one-way sensitivity analysis can be generalized to a 2-way, 3-way, and so on up to an k-way (or all-way) analysis, although graphical representations become prohibitive. A k-way analysis considers all k imprecise quantities at the same time, effectively consider all possible combinations of resolutions of the imprecision. This is important because a one-way analysis ignores interactions between parameters as well as the additive effects. For example, the tornado plot considered only one parameter at its extremes at once. What if actually two of the parameters are at their maxima? This could lead to a combined effect that makes the decision sensitive to the imprecision, a sensitivity that is missed with one-way analysis. In this next section, it is shown that PBA is a generalized and complete all- k-way sensitivity analysis.

6.3 PBA as a generalized sensitivity analysis for decision robustness

By modeling and propagating uncertainty using p-boxes and PBA, a DM can achieve an all-way sensitivity analysis for decision robustness (identifying whether the decision is sensitive to the existing imprecision). It is in this area that PBA is most promising as a sensitivity analysis tool. A discussion of sensitivity analysis for information prioritization (identifying which imprecision is most important to reduce), an area in which the value of PBA is more difficult to assess, is deferred until Section 6.4.3.

It is important to note that there are many ways to compute with p-boxes, just as there are many algorithms for performing an optimization. Therefore, when discussing properties of PBA, it is necessary to specify the algorithms being used. In this section, the set of methods called dependency bounds convolutions (Williamson and Downs 1990, Berleant 1993, Ferson and Donald 1998) (discussed in Section 4.5.1) is considered, and the conclusions drawn about PBA assume the use of these methods except where specifically noted.

The use of PBA as a sensitivity analysis method is not an entirely novel idea. Ferson and co-authors (Ferson, et al. 2004a) explored this question at a high level but without concrete arguments in terms of analysis for decision making and without reference to decision analysis. In this chapter, an argument for PBA as a general sensitivity analysis method is constructed from the bottom up, starting with a simple case. The idea is to show, via a loose type of induction—or more descriptively by informed extrapolation—that the ideas that are easily demonstrable for very simple scenarios extend to more complex scenarios.

For illustrative purposes, the discussion is built up by successively moving to more complicated scenarios. The problem is first limited to two imprecise parameters, $\{{}_1x, {}_2x\}$, so that it can be represented graphically in two dimensions. It is also initially assumed that there are no probabilistic parameters (and hence $k = n = 2$ in the generalized model of Section 3.6). These assumptions will be lifted later in the chapter. As described in Section 6.2.2, a basic sensitivity analysis consists of comparing the range of expected utilities for one alternative with a precisely known expected utility for another alternative. This section will focus on the calculation of this range of performance for one alternative. It is a trivial step to recast the problem for comparisons of alternatives with shared uncertainty discussed in Section 6.2.3

Each imprecise parameter ${}_ix$ is represented as an interval ${}_ix = [{}_i\underline{x}, {}_i\overline{x}]$. Expressed graphically as in Figure 6.3, these ranges correspond to a rectangle (the shaded region), which will be called the consistent region, as it is the region that is consistent with the available information. Also shown are the base values ${}_1\hat{x}$ and ${}_2\hat{x}$ that the DM would use in traditional decision analysis.

If a one-way DASA (as described in Section 6.2) were performed, the dark lines would be explored, probably using some sampling method across this infinite set of points. Notice that these lines represent only a small, and generally non-representative region of the consistent region; considering the entire range of the parameters one at a time (the solid lines) does not yield rigorous or complete analysis of decision robustness.

In order to fully assess the sensitivity of a decision to imprecision, it is necessary to explore the expected utilities of the various alternative actions over the entire consistent region. In general, this would require an all-way sensitivity analysis, which in this simple example is just a two-way analysis.

The goal of sensitivity analysis is to find the interval of expected utility that is consistent with available information. This is equivalent to finding the minimum and maximum of the expected utility in the consistent region—an optimization-like problem. One-way DASA only considers points on the lines, so if the min and max are not on the line, it will not find the true extreme.

The need to explore the entire region raises the question as to how best discretize or sample this region, a question that will become more and more complicated as the dimensionality of the problem increases. The use of interval arithmetic avoids this

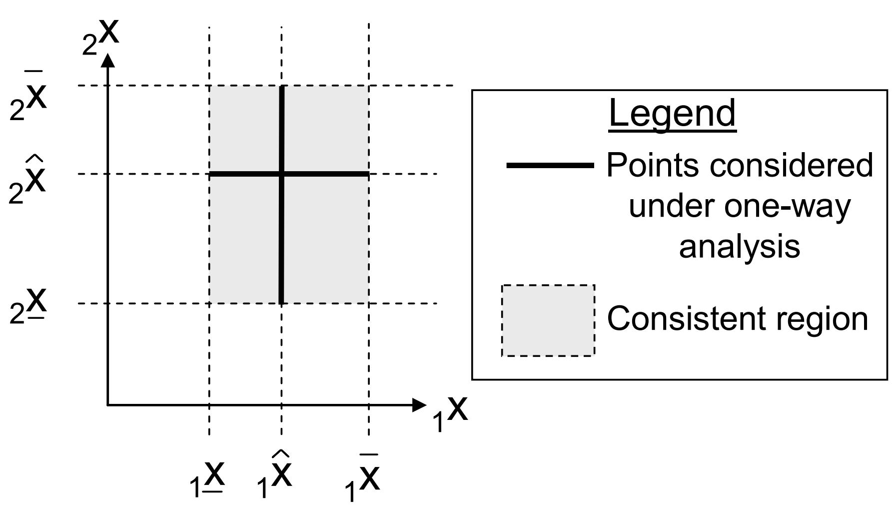


Figure 6.3. Two dimensional imprecise parameter space and sensitivity analysis

problem.

6.3.1 Interval arithmetic as a sensitivity analysis in two dimensions

Since both ${}_1x$ and ${}_2x$ are expressed as intervals, it seems logical to deal with them in their native format and to use interval arithmetic (see for example (Moore 1966, 1979, Kearfott and Kreinovich 1996, Muhanna and Mullen 2004)). Algorithms for interval arithmetic are available that provide rigorous bounds on the output of a calculation. For example, if the true ${}_1x$ is in the interval $[{}_1\underline{x}, {}_1\overline{x}]$ and the true ${}_2x$ is in the interval $[{}_2\underline{x}, {}_2\overline{x}]$, then interval arithmetic methods ensure than for the true result for a function $f({}_1x, {}_2x)$ is contained in the interval that results from the calculation. For example, if ${}_1x = [2,5]$, ${}_2x = [1,7]$, ${}_1\hat{x} = 3$, ${}_2\hat{x} = 5$, and $f({}_1x, {}_2x) = 2 \cdot ({}_1x) + 3 \cdot ({}_2x)$, then $f({}_1x, {}_2x) = [7,31]$.

This example is shown graphically in Figure 6.4 in a format similar to that shown in Figure 6.3. The points in the $\{{}_1x, {}_2x\}$ space that correspond to the minimum and maximum of $f({}_1x, {}_2x)$ are marked with a solid four (✦) and solid five point star (★), respectively. Notice that these points are missed by the one-way analysis in DASA, which finds the interval to be $f({}_1x, {}_2x) = [9,27]$, with the extrema corresponding to the hollow four and five point stars in Figure 6.4. Notice that this interval underestimates the range in both directions (the lower bound is too high, and the upper bound is too low).

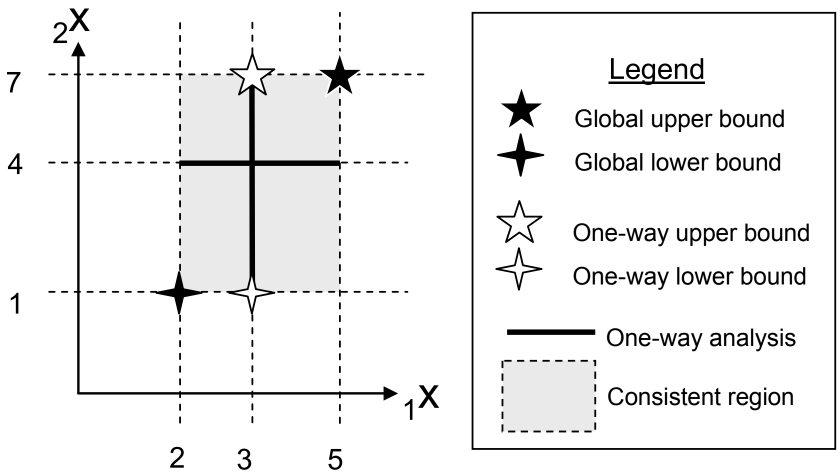


Figure 6.4. Two dimensional imprecise parameter space example problem

The extrema of the output will not always correspond to the corners shown in Figure 6.4. For example, if the function were instead $g({}_1x, {}_2x) = 2 \cdot ({}_1x) - 3 \cdot ({}_2x)$, the maximum of $g({}_1x, {}_2x)$ occurs at (5,1) and the minimum of $g({}_1x, {}_2x)$ at (2,7). In general, the extrema do not have to be the corners at all, so it is actually necessary to search the entire consistent region. Any discretization of this space could miss the true extrema. This was one of the original motivations for developing interval arithmetic: to deliver rigorous results in comparable (or less) computing time than purely numerical methods. Modern methods, such as those summarized in (Muhanna and Mullen 2004), achieve these goals.

6.3.2 Interval arithmetic as a sensitivity analysis higher dimensions

In the preceding section, an example was used to demonstrate how interval arithmetic can serve as a sensitivity analysis. In some ways, this example was too simple to display the real advantage of interval arithmetic as a sensitivity analysis. In this section, the example is expanded into three dimensions, or more specifically, to three imprecise parameters, $\{{}_1x, {}_2x, {}_3x\}$. It is still assumed that there are no probabilistic parameters (and hence $k = n = 3$ in the generalized model). Each imprecise parameter ${}_i x$ is again represented as an interval ${}_i x = [{}_i \underline{x}, {}_i \overline{x}]$.

6.3.2.1 One-way analysis in three dimensions

The consistent region is now a rectangular prism, such as shown in Figure 6.5. Just as in 2-dimensions, the path of points considered in a one-way analysis for one of the imprecise parameters traces out a line. To clarify the figure, the endpoints of these lines are distinguished based on whether they are on the visible faces (solid black circles) of the rectangular prism or on the hidden faces (hollow circles), which also correspond respectively to the upper (${}_i \overline{x}$)and lower (${}_i \underline{x}$) bounds of the imprecise parameters.

In three dimensional space, it is even clearer that the one-way analysis paths do not adequately cover the consistent region, and again the endpoints are even worse. The global minimum and maximum will almost never fall exactly on these lines, and there is no guarantee that the extreme will even be anywhere near these lines. It is necessary to consider more points.

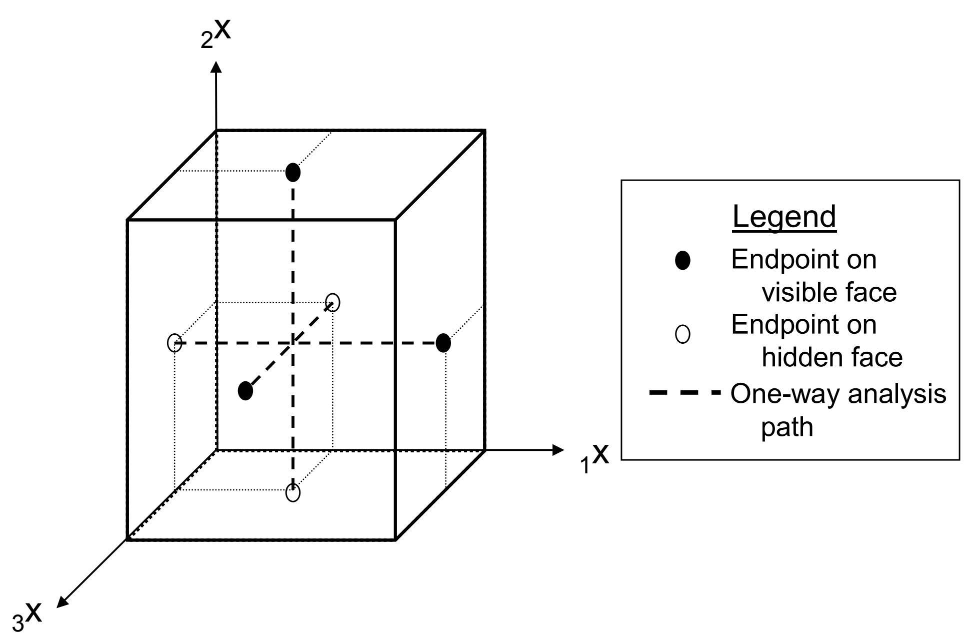
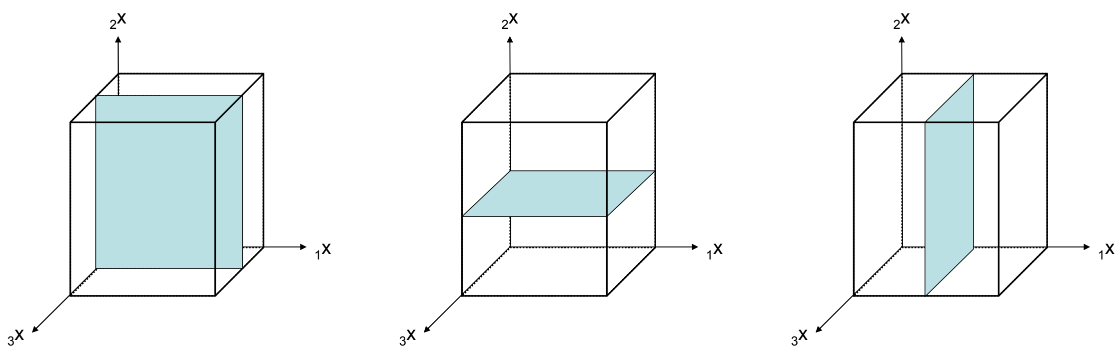


Figure 6.5. Three dimensional imprecise parameter space

6.3.2.2 *Two-way analysis in three dimensions*

Whereas in the two dimensional case a two-way analysis covered the entire consistent region, in three dimensions a two-way analysis only covers part of the consistent space. Since there are three imprecise parameters, there are three possible pairs for a two-way analysis: ${}_1x$ and ${}_2x$, ${}_1x$ and ${}_3x$, and ${}_2x$ and ${}_3x$. The regions that these span are shown in Figure 6.6(a) through Figure 6.6(c) respectively, and then all together in Figure 6.7. These three planes do not cover large regions of the consistent space. This problem will continue in higher dimensions; for a k-dimensional problem, only a k-way sensitivity analysis will cover the entire consistent region, and even then a way must be found to exhaustively search the region. Attention is again turned to the rigorous methods of interval arithmetic.



(a) (b) (c)

Figure 6.6. Planes searched using 2-way sensitivity analysis in three dimensions

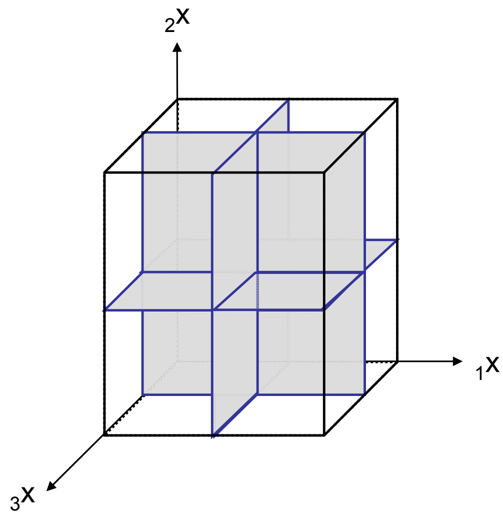


Figure 6.7. Total consistent region searched using two-way sensitivity analysis in three dimensions

6.3.2.3 *Interval arithmetic in three dimensions*

Interval arithmetic becomes more and more useful as the number of imprecise parameters increases, both because it is difficult to visualize more than three dimensions and because thoroughly sampling a multidimensional space is challenging, although methods for searching large combinatorial spaces are available in the design of experiments literature (Neter, et al. 1996, Frey, et al. 2005, Li and Frey 2005) and may be adaptable to this problem. However, interval arithmetic is specifically developed to handle these types of operations.

Interval methods readily extend from two dimensions to three dimensions. This is a trivial result if one accepts binary operations with intervals. For example, even if the function of interest is as complicated as shown in Equation (6.6), it can be decomposed into a sequence of binary operations for which all of the operands are intervals, such as shown in Equation (6.7).

$$f({}_1x, {}_2x, {}_3x) = 2\cdot({}_1x)\cdot({}_2x)^2 + 7\cdot({}_3x)\cdot({}_1x)^2 - 5\cdot({}_2x) \tag{6.6}$$

$$f({}_1x, {}_2x, {}_3x) = \big((2\cdot({}_1x))\cdot({}_2x)\big)\cdot({}_2x) + \big((7\cdot({}_3x))\cdot({}_1x)\big)\cdot({}_1x) - (5\cdot({}_2x)) \tag{6.7}$$

Performing interval calculations such as in Equation (6.7) is equivalent to exhaustively searching the entire consistent prism shown in Figure 6.5. It should be noted that it is generally a bad idea to decompose the equation in this way. By separating the quantities as in Equation (6.7), several quantities appear multiple times, thus exacerbating the repeated variable problem discussed in Section 6.4.1.

6.3.2.4 Interval arithmetic in higher dimensions

Based on the results shown in Equations (6.6) and (6.7), the implementation of interval methods in higher dimensions is trivial. Interval methods are specifically designed to be rigorous in all dimensions, meaning that as long as the true values of the imprecise parameters fall in the hyperspace that describes the consistent input parameters, the true value will fall in the output interval. Because interval arithmetic is equivalent to searching the entire hyperspace of the consistent imprecise parameter values, it is also equivalent to an all-way sensitivity analysis.

6.3.3 Sensitivity analysis with imprecise and probabilistic parameters

In the previous sections, it was assumed that there were no probabilistic parameters. In this section, that assumption is lifted. This scenario is first presented in terms of the traditional DASA approach, and then it is shown how PBA generalizes this approach.

6.3.3.1 Traditional DASA with probabilistic parameters

From one perspective, sensitivity analysis for decision robustness with probabilistic quantities is fundamentally equivalent to sensitivity analysis without probabilistic quantities. Note that the analysis intended in this section is different from probabilistic sensitivity analysis (Homma and Saltelli 1996, Sobol 2001, Chen, et al. 2004, Felli and Hazen 2004, Liu and Chen 2006) mentioned in Section 6.1 which seeks to identify the largest contributors to probabilistic effects in the output. Here the emphasis is on the sensitivity to lack of knowledge (imprecision) about the probabilistic components. However, the presence of a probabilistic quantities does not change the basic process; it only changes the function that needs to be evaluated in the decision problem, since now mathematical expectations need to be calculated over the probability distributions.

In order to handle imprecision about probabilistic parameters in DASA, it is necessary to assume that the type (normal, Weibull, gamma, etc.) of the probability distributions is known. PBA can actually handle the more general case of unknown distributions, as discussed in Section 6.5.1. While DASA assumes that the type of the distribution is known, the parameters of the distributions do not need to be known. For example, consider ${}_9x \sim N({}_1x, {}_2x)$. This means that ${}_9x$ is a probabilistic quantity that is normally distributed with imprecise parameter ${}_1x$ as the mean and imprecise parameter ${}_2x$ as the standard deviation.

When a DASA sensitivity analysis is performed, the imprecise parameters are considered using a single value from their consistent regions at a time. Consequently, for each combination of imprecise parameters (whether one-way, two-way, or more analysis), a particular precise distribution is specified and the expected value can be calculated. This process is then repeated for multiple (discretized) values in the analysis, and the results of these calculations are aggregated into an interval of expected utility.

6.3.3.2 PBA as a sensitivity analysis

In the previous section, the case of an imprecisely characterized probabilistic parameter was considered. The model ${}_9x \sim N({}_1x, {}_2x)$ presented is actually a p-box. Specifically, it is a parameterized p-box. A parameterized p-box is different from a general p-box in that the parameterized p-box consists of a parameterized family of distributions (see Section 4.1). In this case, it is the family of normal distributions with mean ${}_1x$ and standard deviation ${}_2x$. For example, if ${}_1x = [50, 60]$ and ${}_2x = [3, 5]$, then ${}_9x$ can be expressed with the p-box shown in Figure 6.8.

As described in Chapter 4, a p-box is a generalization of both intervals and probabilities. Consequently, it is not necessary to distinguish imprecise and probabilistic variables as separate entity types in PBA as it was in DASA. Some variables, which were previously used only as parameters in probability distributions, can even be eliminated. For example in the above relationship (${}_9x \sim N({}_1x, {}_2x)$, ${}_1x = [50, 60]$ and ${}_2x = [3, 5]$), ${}_1x$ and ${}_2x$ can be eliminated, and the total uncertainty in ${}_9x$ can be written directly as the parameterized p-box ${}_9x \sim N([50, 60], [3, 5])$.

Calculations based on PBA preserve both the intervals (imprecision) and probability information. In general, calculations with p-box inputs result in p-box outputs. Specific algorithms for computing with p-boxes are discussed in Chapter 4. Attention here is limited to a set of algorithms based on dependency bounds convolutions (Williamson and Downs 1990, Berleant 1993, Ferson and Donald 1998). These methods allow for rigorous calculations with p-boxes for the principle binary operators of addition, subtraction, multiplication, and division (as long as zero is not a consistent value). Similar to rigorous interval calculations, a p-box calculations is rigorous if it is true that if the true distributions are contained in the input p-boxes, then the true output

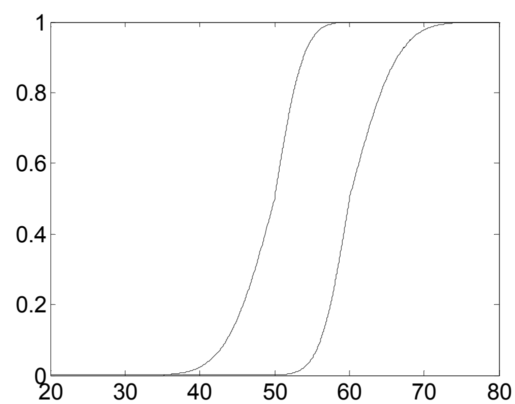


Figure 6.8. Example p-box for ${}_9x$

distribution is included in the output p-box.

Returning to the general mathematical model defined in Section 3.6, the goal of the p-box calculations is to determine the expected utility of particular action. The final output p-box is a set of distributions for the utility of a particular action. Each of these distributions has a particular expected value associated with it, and the entire set in the p-box corresponds to an interval of expected values (see Section 4.6.1). This interval contains all of the expected values that are consistent with the available information, considering all possible interactions and all-way sensitivities.

It is in this way that an uncertainty analysis generalizes sensitivity analysis for decision robustness; the output interval of expected utility is guaranteed to include the true interval, which means that if the true interval crosses the critical value[13], the output interval will, too. Thus, any sensitivity of the optimal action to the imprecision will be detected. If the chosen action's output expected utility interval does not cross the critical value and all inputs were rigorously determined, then the optimality of the chosen action is robust to imprecision.

6.4 Limitations and extensions of PBA-based sensitivity analysis

In Section 6.3, the emphasis was on the advantages of PBA as a sensitivity analysis procedure. Like all methods, the use of PBA is limited to certain scenarios and requires certain assumptions to be made. The overall message of this chapter is twofold: (1) the requirements of PBA are in many ways less strict than the requirements of other

[13] It is important to note that the critical value is known. For comparison between alternatives it is zero. Originally in the literature, an uncertain investment was compared to an investment with a deterministic return, so alternatively the critical value can be defined by some known benchmark.

methods; (2) PBA can provide information that is not available with other methods. Nevertheless, it is important to recognize the limitations of PBA as a sensitivity analysis procedure.

6.4.1 Bounds are rigorous but not necessarily best possible

In the previous sections, it was emphasized that the results of PBA computations using dependency bounds convolutions (DBC) (Williamson and Downs 1990, Berleant 1993, Ferson and Donald 1998) are rigorous, meaning that the true interval is contained in the output interval assuming the inputs are correctly defined. Note that it was not said that the output intervals are the true intervals. This is because the output intervals actually can be larger than the true distributions as a consequence of the repeated variable limitation of interval arithmetic (as explained in Section 4.5.2).

Since the bounds are rigorous, an alternative will never appear to be the most preferred when in fact it is not the most preferred. However, if the bounds are much larger than the best possible bounds (a situation referred to as overly conservative), then there may appear to be significant indeterminacy when in fact the real problem may involve none. This could lead to the conclusion that the decision in not robust when in fact it is robust. This is the opposite of the problem faced by less than all-way sensitivity analysis, which ignores dependencies and higher order interactions and can lead to results that are non-rigorous, i.e., that are inconsistent with the truth.

If a selection decision problem is recast as an exercise in hypothesis testing, the types of errors made with the PBA and sensitivity analyses can be discussed in standard statistical terms of Type I and Type II errors (Devore 1995) via an analogy[14]. Returning to the decision model, $A = \{a_1, ..., a_m\}$ is the set of all possible actions. The null hypothesis is that any one of the $a_i \in \{a_2, ..., a_m\}$ is the optimal action. The alternative hypothesis is that a_1 is the optimal action. Formulated this way, the burden of proof is on showing that a_1 is optimal.

A less than full-way DASA sensitivity analysis may underestimate the true imprecision and indicate that there is enough evidence to reject the null hypothesis in favor of the alternative when there really is not sufficient evidence to do so. In this situation, the null hypothesis would be rejected when it is actually true, a Type I error.

Conversely, PBA may overestimate the uncertainty and lead to the failure to reject the null hypothesis when it is actually false, a Type II error. A Type II error is an error in the sense that an opportunity to make a decision is lost; the null hypothesis could have been rejected, but was not. Consequently, a decision maker may waste resources or make an arbitrary decision trying to reduce indeterminacy that does not exist in the actual problem. PBA will not lead to a Type I error assuming the imprecision in the inputs is sufficiently characterized.

Which is preferable, a Type I or Type II error? A Type II error may be preferable in high-risk applications; when the cost of failure is high, one is often more willing to be conservative and spend additional resources to reduce uncertainty further. In other applications, the cost of delaying a decision or collecting more information may exceed any potential benefit from waiting. There is no general answer; the analyst must assess

---

[14] The alignment is not exact, but is illustrative. As such, this hypothesis testing procedure is meant for discussion only; it is not proposed or recommended as an actual approach to engineering design.

the situation and make his or her own choice. However, one can conclude that PBA leads to a more rigorous sensitivity analysis for robustness in that it avoids a Type I error; it will always detect a lack of robustness.

This result can be seen by considering Figure 6.9 and Table 6.1. In Figure 6.9, the possible resultant intervals from DASA and PBA (using DBC) are compared for (and against) a given truth. These intervals are compared to a set of possible scenarios for the critical value that indicated sensitivity. As shown in Table 6.1, scenarios c_1, c_1', c_2, and c_2' represent cases in which the decision is *not* sensitive. Scenarios c_3, c_3', c_4, and c_4' represent cases in which the decision *is* sensitive. For example, in a scenario in which c_1 represents the critical line, the decision is insensitive to the imprecision (as revealed by the true interval). In this scenario, both DASA and PBA (with DBC) conclude correctly that there is no sensitivity.

These comparisons are summarized in Table 6.1. Clearly whenever the decision is truly sensitive, PBA (with DBC) recognizes this sensitivity (no Type I error) when the inputs are defined correctly. Similarly, whenever the decision is truly insensitive, the DASA approach concludes correctly that it is insensitive (no Type II error). It is also worthwhile noting that since the DASA intervals are always no bigger than the PBA (with DBC) intervals, PBA will conclude that there is sensitivity whenever DASA reaches that conclusion.

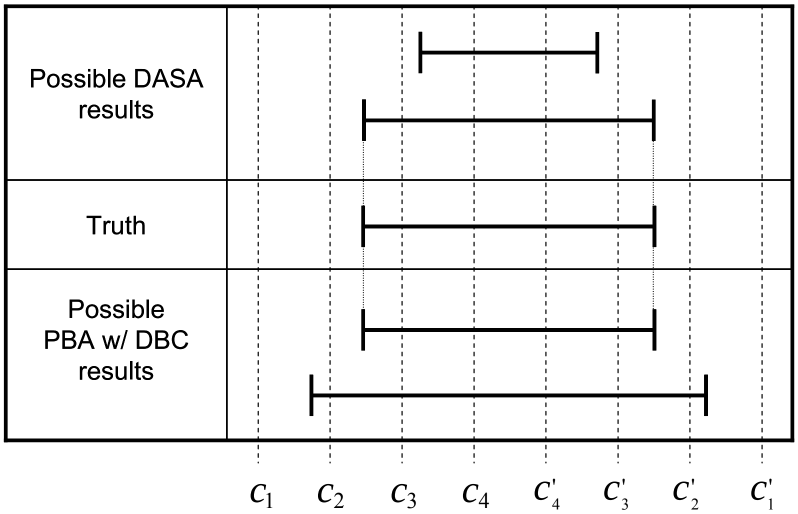


Figure 6.9. Graphical scenarios of sensitivity analysis

Table 6.1. Comparison of sensitivity analysis scenarios, assuming correct inputs

Critical level	Truth	PBA conclusion	DASA conclusion
c_1 or c'_1	Not sensitive	Not sensitive	Not sensitive
c_2 or c'_2	Not sensitive	Sensitive	Not sensitive
c_3 or c'_3	Sensitive	Sensitive	Not sensitive
c_4 or c'_4	Sensitive	Sensitive	Sensitive

In the preceding discussion, it was noted several times that the rigor of the PBA method depends on the accuracy of the characterization of the inputs. This is a somewhat obvious but crucial limitation. For example, if a DM assumes that a quantity is known precisely when actually significant imprecision exists, then any subsequent analysis will underestimate sensitivity. The characterization of inputs is not a perfect science, so underestimation can occur. Consequently, the "guarantee" of rigor does not apply universally, but only when the inputs are correct.

However, if the same assumptions are made for DASA and PBA, then PBA will conclude sensitivity in all cases that DASA does and possibly in more cases. In other words, if the imprecision in the inputs is underestimated, both methods can yield results that underestimate the imprecision in the output, and thus may fail to identify sensitivity. However, there is no case in which DASA avoids a Type I error and PBA does not.

6.4.2 Lack of rigorous black-box methods

As described in Section 4.5.2, DBC methods for p-box computations require an open, operationally defined model (e.g. algebraic) of the problem. Consequently, they cannot be used to analyze so-called black-box models such as differential equations, simulations, and finite element analysis in which the underlying equations cannot be expressed in the appropriate form. The methods for black-box analysis discussed in (Bruns 2006, Bruns and Paredis 2006, Bruns, et al. 2006) generally sacrifice the guarantee of rigor in the calculations. Optimization-based methods often lead to bounds that are close to being both rigorous and best possible, but no guarantee can be made with these heuristic or sampling based methods. Consequently, the value of PBA for rigorous sensitivity analysis for robustness is limited to non-black box models for which DBC are applicable.

6.4.3 Sensitivity analysis for information prioritization

All of the discussion in Sections 6.2-6.3 focused on what was described as sensitivity analysis for decision robustness. Another major use of sensitivity analysis is for identifying to which uncertainties the decision is most sensitive, assuming a decision is sensitive (lacks robustness) given the existing imprecision. PBA does not provide a direct method for determining to which imprecise quantity the decision is most sensitive. Ferson and co-authors (Ferson, et al. 2004a) have proposed a meta-level sensitivity analysis for identifying where future empirical efforts (or information collection) would be most productive.

The meta-level analysis is similar to an approach in probabilistic sensitivity analysis in which the variance of a parameter is reduced and the resulting reduction in the output variance is measured. Ferson et al. propose similarly “pinching” a p-box, where "pinching" means reducing the uncertainty. A p-box can be pinched in two dimensions—reducing the imprecision or reducing the variance of the probabilistic component. Naturally, one could also pinch both dimensions.

Since the goal of a sensitivity analysis (as defined in this chapter) is to identify the sensitivity of the decision to the current imprecision (i.e. lack of available information), it is only reasonable to pinch the p-boxes with respect to the imprecision. Any underlying and inherent probabilistic variability should clearly be preserved.

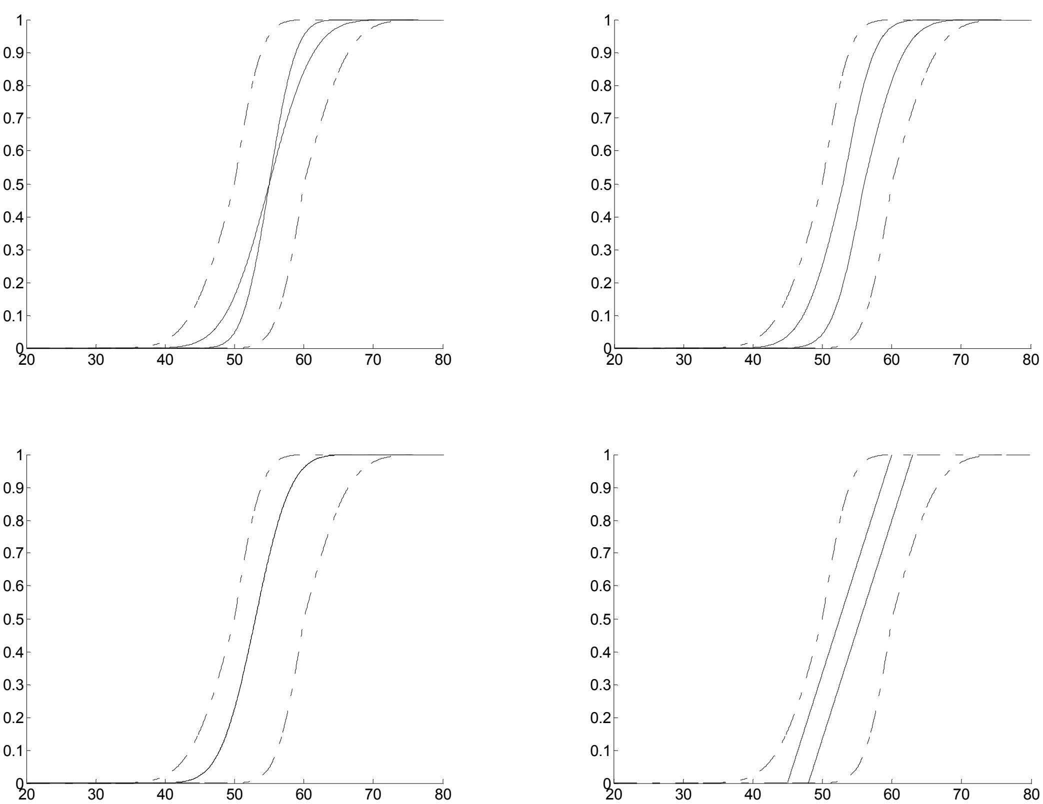
This applies even when subjective probabilities are used. A DM may want to explore the sensitivity of the decision to his beliefs, but if the DM is not sure about his or her subjective probability distributions, he or she should represent this incertitude using a p-box that would contain multiple distributions. The sensitivity of the decision to the imprecision in the DM's beliefs is captured in the p-box analysis.

Regardless of whether the probabilistic information is subjective or objective in nature, a crucial question remains: to which distribution should the p-box be pinched? This question must be answered before the method can be used in practice. There are many possibilities, as shown in Figure 6.10 for the p-box from Figure 6.8. As Ferson and coauthors note, the current answer is pure conjecture (Ferson, et al. 2004a), so future work is necessary.

The answer is not obvious because by its very nature, the p-box contains all of the distributions that are deemed consistent with the available information. One option would be to pinch the p-box to some best-guess distribution (the lower left quadrant of Figure 6.10), such as what would be used in DASA. The idea of this meta-level sensitivity analysis is to pinch a the p-box of a particular input quantity and to compare the resultant p-box using this pinched input to the p-box that results from the original input p-box. The ideal goal would be to identify for which input quantity the pinching reduced the uncertainty in the output p-box the most. This raises another question: how can the uncertainty in two p-boxes be compared?

In probabilistic sensitivity analysis, the difference between two assumptions is measured in uncertainty reduction in terms of variance reduction. Ideally, a meta-level sensitivity analysis also would compare the reduction in uncertainty by pinching a particular uncertain quantity, and then the quantity that yields the greatest reduction in uncertainty would be the priority for additional information collection. Unfortunately, there is no way to fully capture all aspects of the uncertainty in a p-box with a single number, and different measures of uncertainty essentially allow different questions to be addressed (Ferson, et al. 2004a). The validity and exact meaning of such measures of total uncertainty for p-boxes is an open research question.

A DM could perform a qualitative analysis of uncertainty reduction using intuition about the "total uncertainty" in a p-box, but such intuitive methods are less than



(original p-box shown in dotted line, pinched p-boxes shown in solid lines)

Figure 6.10. Pinching a p-box

ideal than formal quantitative methods. However, given the double questions of to what to pinch a p-box and how to measure uncertainty of a p-box, it appears that such intuitive methods will be the best possible in the near future. In that light, the ability of PBA to serve as a sensitivity analysis for information prioritization is currently limited. However, this does not diminish its ability to discern the presence of robustness.

6.5 Additional advantages of PBA as a sensitivity analysis

In addition to providing an all-way, rigorous sensitivity analysis for decision robustness, PBA is also much more flexible that traditional DASA methods. Specifically, PBA allows for unknown distribution types and various conditions of dependency between uncertain parameters.

6.5.1 Unknown distribution types

In Section 6.3, the flexibility of PBA with regard to imprecise distribution parameters was demonstrated, but PBA can also handle cases of unknown distribution type (Ferson and Hajagos 2004). For example, a p-box can be constructed and propagated with only knowledge of the mean and variance; no assumption of distribution type (e.g. normal, lognormal, gamma, or Weibull) is necessary (Ferson 2002). A p-box for a mean of 15 and variance of 20 is shown in Figure 6.11. Such flexibility is useful when, for example, a DM has an estimate of the mean and variance of a probabilistic parameter, but no theoretical or empirical evidence about the distribution family.

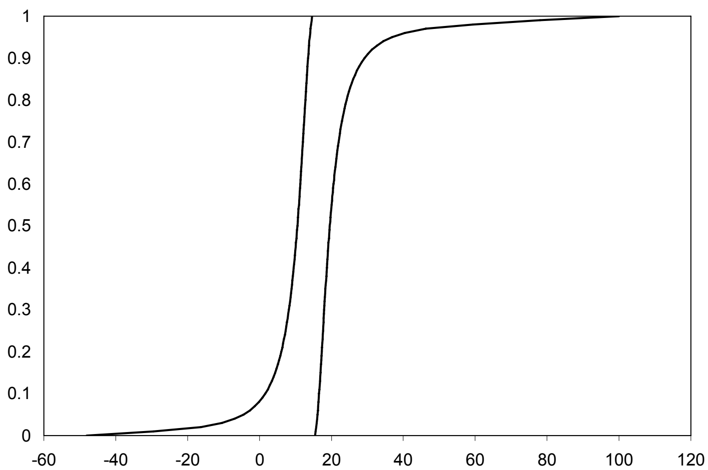


Figure 6.11. Example p-box for mean and variance, no distribution knowledge

6.5.2 Known or unknown dependencies

In many statistical analyses, DMs assume independence between the uncertain quantities. In some cases this is justifiable, but in many cases it is merely a short-cut taken to reduce the complexity of the problem. The challenge is that traditional statistical methods require that the joint probability density functions between all uncertain quantities to be known precisely. Since information regarding joint density functions is often scarce, it is convenient to assume independence, thus making the joint density functions simple functions of the marginal density functions.

DBC methods can determine probability bounds in the case of unknown dependence between the inputs. Essentially, a p-box can be computed that covers all possible dependency scenarios; this allows the DM to avoid making unwarranted assumptions about dependencies. DASA sensitivity analysis, on the other hand, ignores dependencies and higher order interactions, and it requires a known distribution type. Consequently, the class of problems that can be accurately explored with PBA (using DBC) is more inclusive that the class of problems that sensitivity analysis can explore.

6.6 Summary: General usefulness of PBA

In this chapter, it was shown that PBA generalizes a global (or all-way) sensitivity analysis, as defined in decision analysis. When implemented using the DBC algorithm, PBA provides a means for performing a sensitivity analysis for robustness identification that avoids a Type I error; it will never lead to the conclusion that the decision is not sensitive to the existing lack of information when in fact the decision is sensitive, assuming the imprecision in the inputs are properly characterized. In other words, it never leads to a false sense of security. PBA can also handle a wider array of problems that a traditional sensitivity analysis because it can be used to propagate uncertainties without making any assumptions about the dependencies between uncertain quantities.

However, PBA also has important limitations. First, it provides no direct information for prioritizing information collection; it identifies the existence of sensitivity, but not the source. Second, PBA is subject to a Type II error, which means concluding that the decision is sensitive to the existing lack of information when in fact it is not sensitive. This could result in missed opportunities or unnecessary expenditures on information collection. The exact tradeoffs between the two methods of sensitivity analysis are explored in more detail in the next chapter, which compares the two for the specific example of the environmentally benign design of an automotive oil filter.

CHAPTER 7:

DEMONSTRATION OF PBA AND A COMPARISON TO ONE-WAY SENSITIVITY ANALYSIS IN THE CONTEXT OF ENVIRONMENTALLY BENIGN DESIGN AND MANUFACTURE

This chapter, much of which was previously published in a conference paper (Aughenbaugh, et al. 2006), serves as a capstone to the core of the dissertation. In Chapters 2 and 3, motivating questions 1 and 2 were explored and potential answers were advanced. In Chapter 4, the details of a particular representation were presented. Chapter 5 contained a motivating example that demonstrated the potential value, in terms of the quality of the final design, of using imprecise probabilities. In Chapter 6, it was shown how PBA presents a rigorous sensitivity analysis for robustness, as compared to decision analysis with sensitivity analysis (DASA). In this chapter, aspects from each of these chapters are brought together and augmented in the context of the environmentally benign design of an oil filter. The goals of the chapter are to:

1. Explore an application in which the ability to represent imprecision is important and could have significant impact on society
2. Demonstrate the process of performing a probability bounds analysis
3. Compare PBA and decision analysis with one-way sensitivity analysis in a specific context

The first section of this chapter is an introduction to the context of environmentally benign design and manufacture. The example problem of oil filter selection is introduced in Section 7.2. The process of applying PBA to the selection problem is demonstrated in Section 7.3. The decision analysis with one-way sensitivity analysis approach to solving the same problem is presented in Section 1.4. In Section 1.5, the two methods are compared in the context of the oil filter example.

7.1 Environmentally benign design and manufacture (EBDM)

Consumers and legislators are beginning to recognize the cost to society that results from the environmental impact of products. Consequently, companies are increasingly concerned with the environment. Interest is therefore growing in environmentally benign design and manufacture (EBDM), a domain that examines the often competing goals of achieving economic growth and protecting the environment.

The economic goals of product design are relatively clear and generally can be tracked using financial accounting. Environmental goals are harder to quantify. It is challenging to assess the environmental impact of a product because there are many people who can suffer from environmental impacts, and these people are often spread out geographically and temporally. For example, in addition to manufacturing workers and customers who are in direct contact with the product, the residents in the areas where products are produced and disposed of can feel the effects of the products. Residents in areas downstream and downwind from industrial sites can also suffer.

These events can take place at many different locations in many different neighborhoods (and even countries) over the product's entire (and often long) lifespan and even long into the future. Consequently, an evaluation of all of the loads and impacts has traditionally been addressed with life cycle assessment (LCA) methods. Researchers are starting to recognize that a key characteristic of LCA is that only very limited information and knowledge is available, resulting in large uncertainty, as summarized by Ross and coauthors (Ross, et al. 2002) and Björklund (2002).

In general, multi-criteria evaluations that include environmental performance can be decomposed as depicted in Figure 7.1. Similar decompositions have been proposed

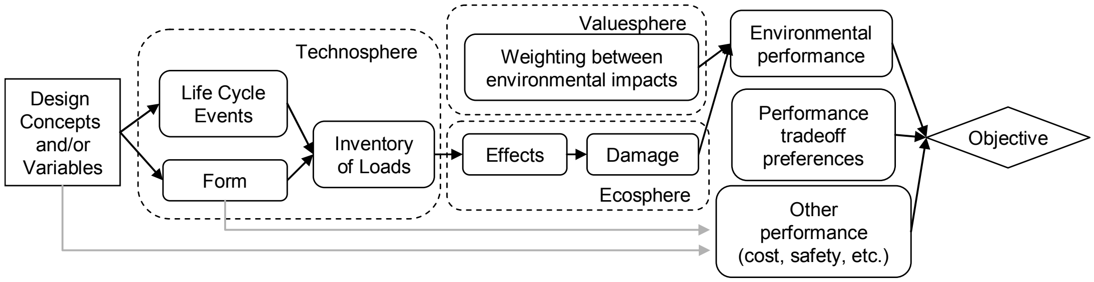


Figure 7.1. The components of an environmental analysis

(Hofstetter 1998, Lu and Gu 2003), though none are identical in form or scope to the structure presented here. Components are grouped, as indicated by dashed-lines in the figure, using Hofstetter's concept of "spheres" of knowledge and reasoning about environmental evaluation (1998).

- Technosphere: description of the product and its life cycle and an inventory of loads (e.g. emissions)
- Ecosphere: modeling of changes to the environment
- Valuesphere: modeling of the perceived seriousness or importance of changes to the environment

Any of the components in Figure 7.1 can be a source of uncertainty. Often some of these sources, such as form and inventory, are well characterized, while others, such as environmental damages, are much harder to characterize.

EBDM is a multi-objective problem, pursuing the often competing goals of economic growth and environmental protection, while subjected to multiple sources of uncertainty. This is a rich context in which to explore different methods of representing uncertainty and making engineering design decisions. It also offers an opportunity to contribute to the EBDM and LCA communities by demonstrating practical approaches for uncertainty management in those domains.

7.2 EBDM Example: Selecting an appropriate oil filter design

Around 250 million light duty oil filters are discarded (and not recycled) in the United States each year. The environmental impact of these filters can be substantial, as disposable filters contain large amounts of steel, aluminum, or plastic, depending on the style of filter.

In this example, it is assumed that an automobile manufacturer wants to reduce the environmental impact of oil filters from its cars by designing a more environmentally benign filter. Naturally, the company simultaneously wants to turn a profit, making this an EBDM problem. It is assumed that since high-price filters are less attractive to consumers than low-price filters (with all other things being equal), the manufacturer wants to minimize the total cost to the consumer of purchasing oil filters over the lifetime of the vehicle.

In the following, the example problem is described in detail. This extensive explanation will be useful in the subsequent sections, as a complete understanding of the example problems and assumptions is necessary to appreciate the advantages and limitations of the methods.

Naturally, some simplifications and assumptions are introduced in the problem. For example, the exact dimensions and parameters for the problem are chosen to be realistic, but are not based on hard, real-world data. Consequently, the emphasis is not on the actual decision outcome (i.e. the chosen filter), but rather on the decision and analysis process.

7.2.1 Types of oil filters

In this simplified model, shown in Figure 7.2, an oil filter is comprised of five components: housing, top cap, filter, inner support, and bottom cap. The housing, top cap, and bottom cap make up the casing, and the inner support and filter make up the cartridge. Three different types of oil filters are considered, as summarized in Table 7.1.

For the steel easy change (SEC) filter, the structural components are made of steel. The entire filter is designed to be replaced at once; it is simply unscrewed from the engine and then discarded or recycled. The plastic easy change (PEC) filter is used exactly as the SEC filter, but its structural components are plastic rather than steel. Finally, the take-apart spin-on (TASO) filter has structural components made of aluminum and when the filter is replaced, only the cartridge is replaced; the casing is reused.

7.2.2 The design problem

The design problem used in this example is a selection between the SEC, PEC, and TASO filter types. The "best" choice depends on various factors, as shown using an

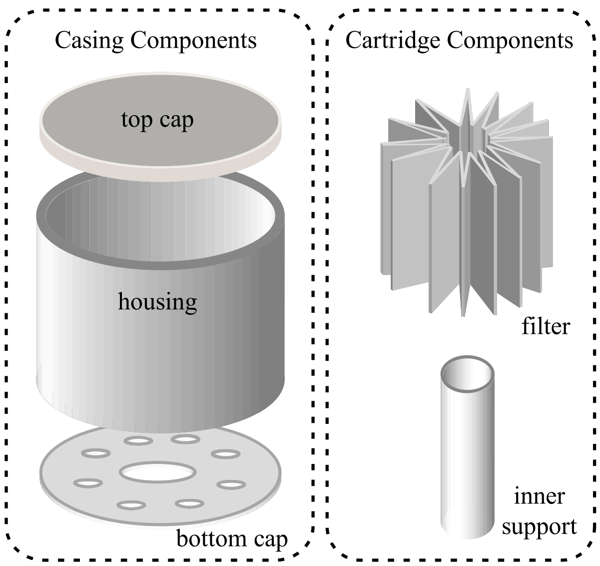


Figure 7.2. Oil filter schematic diagram

Table 7.1. Types of filters

Filter	Material	Discarded parts
SEC	Steel	Cartridge and Casing
PEC	Plastic	Cartridge and Casing
TASO	Aluminum	Cartridge only

influence diagram in Figure 7.3 [see (Clemen 1996) for an introduction to the usage and (McGovern, et al. 1993) for an overview of the history of influence diagrams]. The influence diagram is constructed by first identifying the decision (i.e. filter type selection) the objective function; the purpose of the influence diagram is to map what influences this objective.

7.2.3 Objective function

It was noted earlier that the manufacturer has two primary goals: to minimize environmental impact, and to keep the cost to the user low. This is shown in Figure 7.3 by the arrows leading from "Total user cost" and "Total impact on environment" into the objective. In this example, the objective will be expressed in terms of utility, using multi-attribute utility theory (Keeney and Raiffa 1993, Scott 2004).

Specifically, the total utility is given in Equation (7.1), and the individual utility functions are given by Equations (7.2) and (7.3).

$$U_{total} = U_{cost} + w \cdot U_{impact} \tag{7.1}$$

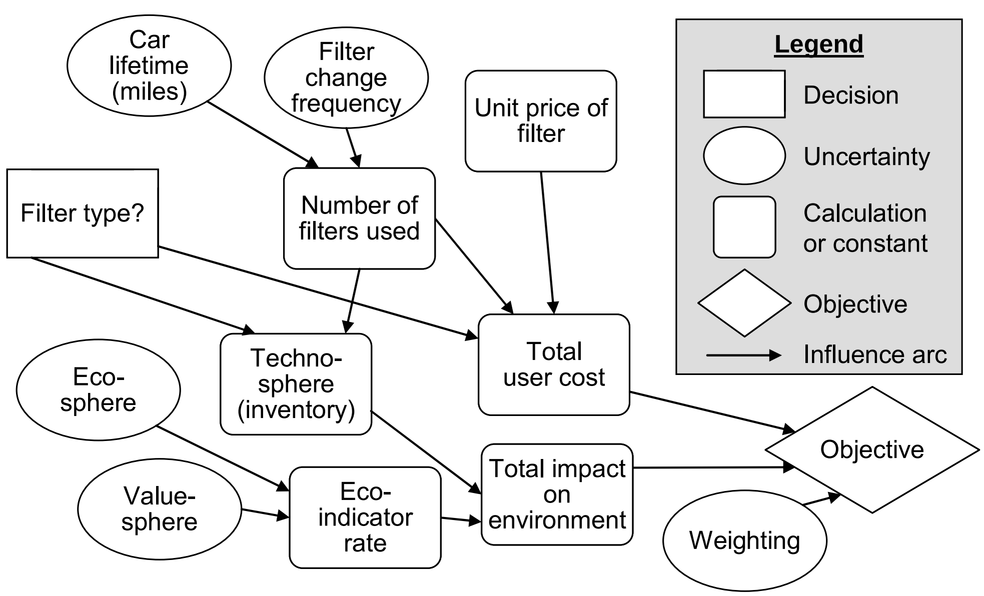


Figure 7.3. Influence diagram of the decision problem

$$U_{cost} = -cost \tag{7.2}$$

$$U_{impact} = -ecoscore \tag{7.3}$$

The quantity ecoscore is explained in the next subsection. As defined, all utilities will be negative, but they remain relevant for comparisons. This simple utility function is chosen to highlight the influence of uncertainty representations on the decision process. The elicitation of exact multi-attribute utilities is challenging, so the weight parameter w is assumed to be imprecise, confined only to the interval $w = [0.5,\ 2]$, with the best guess at $w = 1$. In practice, the functional form of the utility functions could also be uncertain, a scenario which is not addressed in this dissertation.

7.2.4 Environmental impact calculation

It is assumed that the primary environmental impact of an oil filter arises due to the construction, transportation, and disposal of the casing and cartridge. These components are constructed of materials such as steel, aluminum, and plastics and are present in large quantities. Other substances, such as the cellulose filter element and oil residue, are present in much smaller quantities and are generally equivalent in all three types of filters.

The Eco-indicator 95 is an impact assessment method for life-cycle analysis in which particular scores (ecoscores), measured in eco-points, are assigned to specific materials and processes There is also an updated Eco-indicator 99 available (Goedkoop and Spriensma 2001), but for illustrative purposes the old database and methodology is sufficient. Assuming that the impact of geometric forming processes is negligible compared to the impact of the mass of the material and its production (mining, processing, etc.), these scores can be given for specific materials as points per mass. They will be referred to as Eco-indicator rates, or simply *ecorates* in this dissertation. The impact of a filter of material m is computed using Equation (7.4).

$$impact_m = ecoscore_m = mass_m \cdot ecorate_m \tag{7.4}$$

For a particular material, the ecorate not only captures its environmental effects and damages, but also sets a value on these damages relative to other damages. As such, it combines the ecosphere and Valuesphere components of Figure 7.1. This allows for tradeoffs between different materials and processes with different inherent environmental impacts. These value tradeoffs are fixed within the Eco-indicator model, but in practice, not every society or decision maker will agree with these tradeoffs. Consequently, the Valuesphere is a source of uncertainty in environmental life-cycle assessment. The effects and damages are uncertain due to the complexity and uncertainty of modeling ecosystems.

In Figure 7.1 the ecosphere is independent of the Technosphere because the ecorates are independent of what effects are present; they are pre-tabulated for all materials. In this problem, the Technosphere effects, or inventory, is given by the total mass of material used over a vehicle's lifetime, and are thus incorporated into the problem in Equation (7.4).

The total mass depends on the number of filters F used, as shown in Table 7.2. This number varies because not every vehicle is in service for the same number of miles, and car owners change the filters with difference frequencies. Both vehicle lifetime L and frequency of filter change f of a vehicle are thus uncertain. The number of filters used over a vehicle's lifetime is given by Equation (7.5) as a function of these two uncertain quantities, which were also reflected in Figure 7.3.

$$F = L / f \tag{7.5}$$

Table 7.2. Total environmental impact and cost functions

Filter	Total mass of F filters (kg)	Total cost of F filters
SEC	$\left(mass_{casing} + mass_{catridge}\right) \cdot F$	$\$5 \cdot F$
TASO	$mass_{casing} + \left(mass_{catridge}\right) \cdot F$	$\$15 + \$5 \cdot F$
PEC	$\left(mass_{casing} + mass_{catridge}\right) \cdot F$	$\$8.5 \cdot F$

7.2.5 Total user cost calculation

The total cost is a function of the number of filters F used over the vehicles lifetime and the price of filter parts. This function depends on the type of filter, since different parts are used for different filter types. The functions and prices are summarized in Table 7.2. In general, filter price is uncertain due to unknown market fluctuations and inflation. For simplicity in this example, the costs are assumed known and constant.

7.2.6 Assumptions on available information

The assumed uncertainties are summarized in Table 7.3. The environmental impact per unit mass is known to be within a stated interval for each material. Interval data is assumed because the uncertainty in the numbers is not probabilistic but rather arises from modeling errors and assumptions about the ecosphere and Valuesphere. The stated intervals represent 50% to 200% of the nominal values calculated using the Eco-indicator 95 analysis (6.2, 21.2, and 9.4 millipoints per kilogram for SEC, TASO, and PEC filters respectively).

Table 7.3. Assumptions about uncertainty

Uncertain parameter	Assumption	
Vehicle lifetime L (miles)	$Gamma(\alpha_1, 2)$ $\alpha_1 = [40000, 60000]$	
Filter change frequency f (miles/filter)	$3000 + Weibull(\alpha_2, 5)$ α_2=[3000, 5000]	
Eco-impact rate (millipoints/kg)	Steel	[3.1, 12.4]
	Plastic	[10.6, 42.4]
	Aluminum	[4.7, 18.8]
Utility weighting w	[0.5, 2.0]	

Imprecise probabilities are used to represent the uncertainty in the vehicle lifetime and filter change frequencies. Variability (a type of irreducible uncertainty) arises because the population of vehicle owners contains a variety of individuals, each of whom has his or her own behavior, but who collectively appear random. For illustration, only one parameter of the distributions is assumed to be known imprecisely, but the methods immediately generalize to multiple uncertain quantities.

It is assumed that L and f are both independent of all three eco-rates, and that the weighting w in the utility function is independent of all other quantities. However, the dependency between L and f is unknown, as are all dependencies between all ecorates.

Why is independence not assumed? L and f are both related to user behavior. It is conceivable that a user who intends to keep a car a long time will change the filter at a higher rate than someone who keeps a car a short time, since the long-time owner would have a greater interest in keeping the engine in good condition. In such a scenario, L and f are correlated. However, this dependence is not known exactly and may not even exist at all, so it makes sense to assume an unknown dependency. A similar argument can be made between the eco-rates; they could be independent since they relate to different material and potentially different environmental effects and damages. However, they could also be correlated if, for example, they share an effect in the ecosphere.

A traditional statistical approach would require perfect knowledge of all joint probabilities, information that is rarely known. Consequently, independence between uncertain quantities is commonly assumed, an assumption that is often unjustifiable given available information. The ability of PBA to handle unknown dependencies, and therefore compute the possible range of results with just the marginal distributions as inputs, is a major advantage over traditional methods.

7.3 Oil filter selection using PBA

The objective function in this example, Equation (7.1), has two components: cost and environmental impact. Of the two, the cost calculation facilitates a clearer demonstration of the PBA process. Consequently, the single objective of cost is considered separately from environmental impact in order to illustrate some properties of PBA. Attention is subsequently returned to multi-attribute EBDM problem in Section 7.3.6. The commercially available Risk Calc (Ferson 2002) software is used for all calculations in this section. Risk Calc implements PBA using the DBC algorithms discussed in Section 4.5.1.

7.3.1 Total cost calculation

According to Figure 7.3, total cost is a function of the number of filter changes F, which according to Equation (7.5) is strictly the quotient of L divided by f. This calculation is evaluated in Risk Calc by first defining the variables L and f according to Table 7.3, which results in the p-boxes shown in Figure 7.4 and Figure 7.5 respectively. The result of the quotient is then calculated directly using Equation (7.5). This results in the p-box for the total number of filter changes F shown in Figure 7.6.

To calculate the total cost, the p-box is merely scaled and shifted according to the equations in Table 7.2. The result for the SEC filter is shown in Figure 7.7. The p-box in Figure 7.7 captures the set of cumulative probability distributions for total cost of the

SEC filter that are consistent with the available data. The procedure can be repeated for the TASO and PEC filters using the appropriate equations. The next step is to compare the results for the three alternatives.

7.3.2 Comparing p-boxes

In traditional statistical decision theory, precise probability distributions are assumed, and a decision maker compares alternatives by taking mathematical expectation over the distributions and selects the alternative with the lowest expected cost. For a p-box, a similar calculation can be made, but since the p-box consists of multiple distributions, multiple expected costs result, which together form an interval of expected cost for each alternative (as discussed in Section 4.6.1).

The intervals of expected cost can be calculated using the bounding distributions

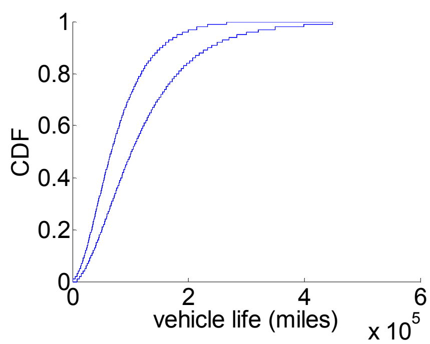


Figure 7.4. Probability box for vehicle life

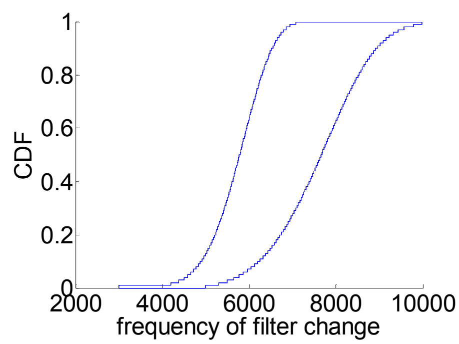


Figure 7.5. Probability box for filter change frequency

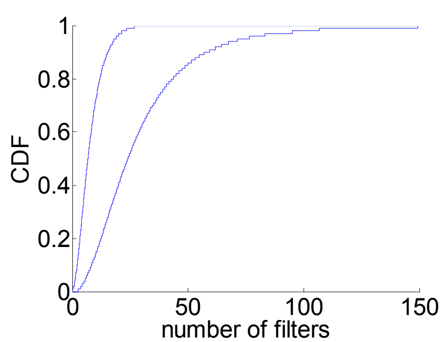


Figure 7.6. Probability box for total number of filter changes

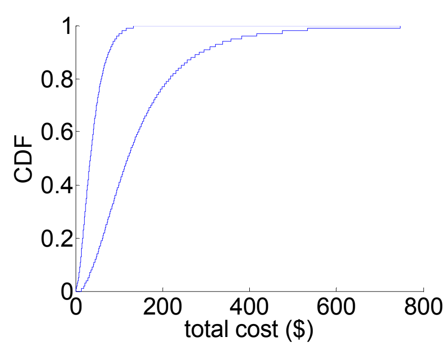


Figure 7.7. Probability box for total cost of SEC filter

of the p-box (see Section 4.6.2). The resultant intervals of expected cost are shown in Figure 7.8. When two intervals of expected cost overlap, as in Figure 7.8, the choice of the best alternative is indeterminate. Depending on where in the two intervals the true values lie, either one could be the best choice, but the available information does not present a rational way to determine which really is. In some cases, this indeterminacy can be resolved, as follows.

7.3.3 Resolving indeterminacy

When the basic analysis results in indeterminacy of choice, the decision maker has several options, including:

1. Collect additional information
2. Make an arbitrary, satisficing, or robust decision
3. Base the decision on additional criteria
4. Include information in the decision that is not captured in the intervals, such as shared uncertainty

The first option is straightforward—collect information until the indeterminacy is eliminated or until it is no longer valuable to reduce it. As an example of the second,

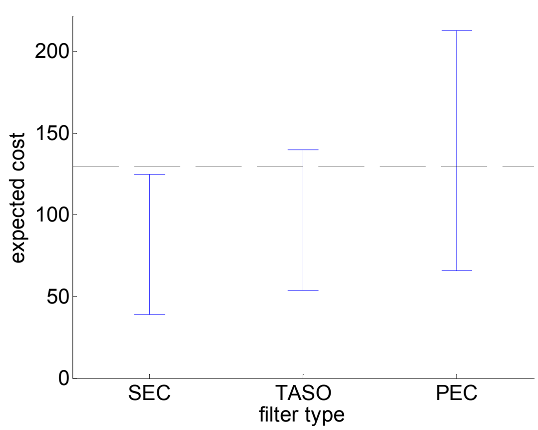


Figure 7.8. Intervals of expected cost

assume that the decision maker is convinced that an oil filter with a lifetime cost to the average user of greater than $130 is unacceptable. In that case, any filter whose interval is above or includes this value is unacceptable. In a satisficing approach, any alternative that meets this constraint is acceptable. In this example, only the SEC filter meets this criterion with certainty, so it would be chosen. According to the intervals in Figure 7.8, it is possible that the PEC and TASO filters perform better than the SEC filter, but the SEC filter is guaranteed to meet the minimum requirement, while the other are not.

One possible arbitrary policy is to choose the alternative with the lowest lower bound (a so-called mini-min policy), which suggests risk-taking or optimistic behavior. An opposite policy is to choose the alternative with the lowest upper bound (a so-called mini-max policy[15]), which suggests risk-aversion or pessimism. In this example, all three policies result in the same choice, but this will not always be the case.

The third option is to include additional criteria into the decision. In this example, the decision maker's actual goal is to minimize both the cost and environmental impact of the filter, so the impact information could be included in the decision with the hope of resolving the indeterminacy.

The fourth option for resolving indeterminacy is to include additional information. This is possible because any mathematical representation of information is an abstraction; generally, some information is lost when it is forced into a particular formal representation. It is not always economical to use the most expressive formalism, because the complexity of a formalism increases the computational and informational

---

[15] The "maxi-min" policy is referred to more frequently in the literature, but the policies are essentially the same; maxi-min is used in maximization problems, and mini-max is used when the objective is a minimization problem.

costs. In the filter example, one possibility source of additional information to include is shared uncertainty (Rekuc, et al. 2006).

7.3.4 Considering shared uncertainty

Accounting for shared uncertainty when handling imprecision is similar in goal to using joint probability distributions or correlations when handling precise probabilities; the goal is to explicitly recognize events that tend to occur together (or separately). In this example, the number of filters F used is assumed to be independent of the filter design; the choice of filter has no impact on the number of times it will be replaced. Consequently, the imprecision in F is the same, or shared, for all three alternatives.

In the preceding analysis, this commonality was not considered. For example, the upper bounds on the intervals shown in Figure 7.8 actually all correspond to the distribution of F with the largest mean value. Consequently, it only makes sense to compare the upper bound of the SEC cost to the upper bounds for TASO and PEC rather than the entire intervals. One way to account for such shared uncertainty is to compare the differences between the alternatives at every realization of F, a so-called maximality comparison (Walley 1991, Rekuc, et al. 2006) [see also Chapter 4] that has a similar motivation as does paired statistical testing (Devore 1995). The need for considering these differences explicitly is further explained in the following section, but first the results are presented.

The resulting cost difference intervals are shown in Figure 7.9. In this figure, the intervals are not compared directly to each other, but rather each interval is compared to zero. Starting at the left, the interval for the expected difference in cost between the SEC filter and the PEC filter ($cost_{SEC} - cost_{PEC}$) lies entirely below zero, so (since the decision maker wants to minimize cost) the SEC filter is preferred to the PEC filter (based on expected cost alone). Similarly, the interval for the difference between SEC and TASO is entirely below zero, meaning SEC is preferred to TASO. Consequently, it becomes clear

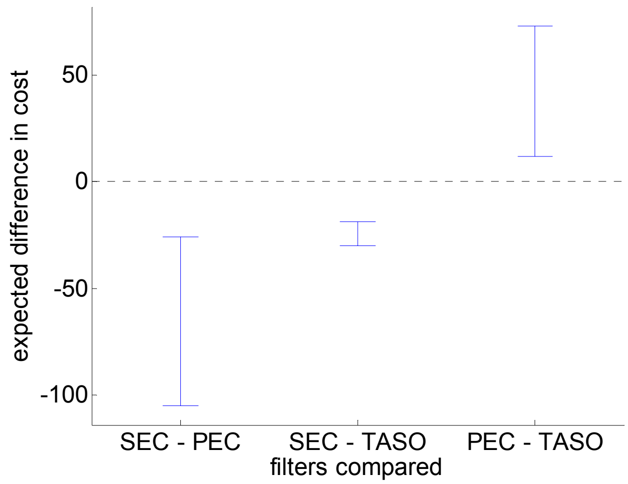


Figure 7.9. Intervals for expected difference and cost

that SEC is the best alternative in terms of minimizing expected cost, a conclusion that could not be drawn from Figure 7.8 when shared uncertainty was ignored. The issue of shared uncertainty is addressed in detail in Chapter 8.

7.3.5 Interval arithmetic and repeated variables

In the previous section, the importance of considering shared uncertainty when comparing alternatives was illustrated. The explicit consideration of shared uncertainty in PBA is necessitated partly by a fundamental limitation of interval arithmetic. As described in Section 4.5.1, interval arithmetic algorithms are in general rigorous but not best possible.

Since the bounds are rigorous, an alternative will never appear to be the most preferred when in fact it is not the most preferred. However, if the bounds are much larger than the best possible bounds (a situation referred to as overly conservative), then there may appear to be significant indeterminacy when in fact the real problem may involved none.

For example, the p-boxes for differences in utility (both considering and not considering the shared uncertainty) between the SEC filter and TASO filter are shown in Figure 7.10. The true p-box (solid line) is actually much more restrictive than the one (dotted line) calculated ignoring the presence of repeated variables. An overly conservative p-box will lead to intervals of expected utility that are larger than necessary, possibly causing unwarranted indeterminacy in the decision, as described in Section 6.4.1. The consequence of this for the oil filter example is discussed in Section 7.5.2.

7.3.6 Multiple objective analysis and selection

In the previous section, various properties of PBA were demonstrated and discussed in the context of cost minimization. Attention is now returned to the actual multi-attribute example problem. Using Risk Calc, all of the uncertainties shown in Table 7.3 can be propagated through the problem to evaluate overall utility. The resulting intervals for the comparisons of design alternatives are shown in Figure 7.11. Based on these results, no decision can be made because all three intervals contain zero. Such indeterminacy can be expected if the uncertainty is large, because large uncertainty implies a lack of information for determining the best alternative.

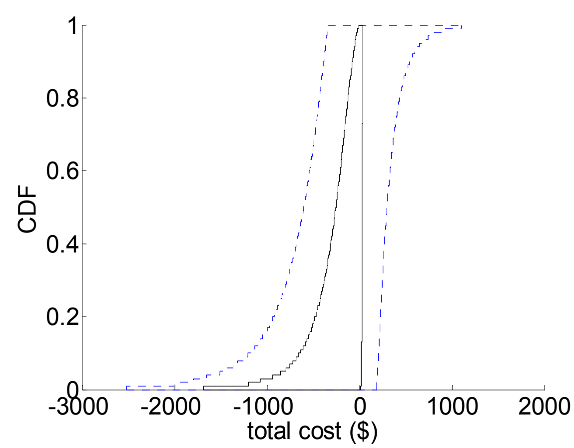


Figure 7.10. Cost p-boxes for the quantity (SEC minus TASO)

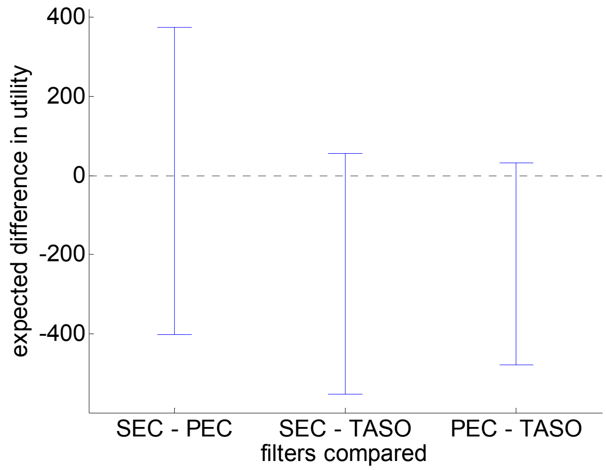


Figure 7.11. Intervals of expected difference in utility

The indeterminacy implies that more information or a different decision policy is needed, such as discussed in Section 7.3.3. However, because the PBA calculations are not always best possible, the indeterminacy shown in Figure 7.11 may not actually exist. This point is revisited in the discussion Section 7.5.1.

7.4 Oil filter selection with decision analysis and sensitivity analysis

Decision analysis is a discipline that studies procedures, tools, and frameworks for transforming problems that are difficult to understand, solve, or explain into problems that are more readily understood and solved (Howard 1988a, Clemen 1996). This section presents a basic decision analysis, including sensitivity analysis, approach to selecting an oil filter.

7.4.1 Basic decision analysis

Clemen (page 6) describes the decision-analysis process with the following steps:

1. Identify the situation and understand objectives
2. Identify alternatives

3. Decompose and model the problem structure, uncertainties, and preferences
4. Choose the best alternative
5. Perform sensitivity analysis
6. Decide if further analysis is needed
7. Implement the chosen alternative

This process will be referred to as decision analysis with sensitivity analysis (DASA) in this chapter. Steps 1-3 have been completed in Section 7.2. The resulting data is the same as in Table 7.3, but the uncertainties will be treated differently. Traditional decision analysis does not explicitly recognize imprecision, so the analysis in step 4 is performed using best-guess, base values for all of the imprecise parameters, as shown in Table 7.4. Notice that this data is consistent with the data in Table 7.3.

The expected value of the objective function for each design alternative can be calculated using the relationships described earlier. The results are shown in Table 7.5. Assuming the base values for all quantities, the TASO filter has the highest expected utility and is therefore the best alternative (recall that as the utility function was defined in Equations (7.1)-(7.3), all of the utilities were guaranteed to be negative, and now the TASO filter has the least negative expected utility and is therefore the best). The next step in decision-analysis is to perform a sensitivity analysis on the selection of the TASO filter, as discussed in the next subsection.

Table 7.4. Base values for imprecise quantities

Uncertain parameter	Assumption	
Vehicle lifetime L (miles)	$Gamma(50000, 2)$	
Filter change frequency f (miles/filter)	$3000 + Weibull(4000, 5)$	
Eco-impact rate (millipoints per kg)	Steel	6.2
	Aluminum	21.2
	Plastic	9.4
Utility weighting w	1.0	

Table 7.5. Results with nominal values

Filter	Expected utility
SEC	-167
TASO	-119
PEC	-196

7.4.2 Sensitivity analysis

The use of the base values in the previous analysis ignores the knowledge about the uncertainties that was described in Table 7.3. In a sensitivity analysis, a decision maker asks, “How might this neglected uncertainty affect the decision?”

As described in the opening of Chapter 6, the phrase sensitivity analysis has been used to refer to various procedures in engineering design. One convenient way of performing a sensitivity analysis for a selection decision is to evaluate the sensitivity of the decision outcome graphically using tornado diagrams (Howard 1988a, Eschenbach 1992, Clemen 1996), such as shown in Figure 7.12 for the SEC filter. A tornado diagram allows a decision maker to perform a one-way sensitivity analysis—that is, to explore the effects of uncertain quantities one at a time.

The first step in constructing a tornado diagram is to define upper and lower bounds for the uncertain parameters, such as those shown in Table 7.3. The next step is to take one parameter and vary it from its lower limit to its upper limit with all other quantities held constant at their base values. This results in an interval of values for the objective function. Finally, this is repeated for each parameter. The base values are marked with dashed lines. Note that since the SEC file contains no plastic or aluminum, it shows no sensitivity to the ecorates for these materials.

The tornado diagram is useful for determining which quantities have a large potential to affect the decision, given the stated uncertainty bounds. This is often done by comparing the performance to a reference line that represents the performance of a precisely characterized alternative. If a bar crosses this reference line, then the decision is (one-way) sensitive to the corresponding parameter. This information can then guide information collection or modeling decisions.

In the oil filter example, all three alternatives involve uncertainty, so there is no constant reference line for comparison. Instead, the difference in performance between alternatives can be compared for each of the necessary pairings, such as shown in Figure 7.13. For pair-wise comparisons, the reference line becomes zero. If a bar of the tornado plot crosses the zero reference line, it indicates the preferred alternative can switch depending on the outcome of the uncertainty.

According to the results, the preference of TASO over SEC is one-way insensitive to the uncertainties. The preference of TASO over PEC is also one-way insensitive to the uncertainties. Based on a strict interpretation of the one-way sensitivity analysis, the choice of the TASO filter is robust to the uncertainty, a conclusion that contradicts the findings of the PBA analysis. Further comparisons are addressed in Section 7.5.1.

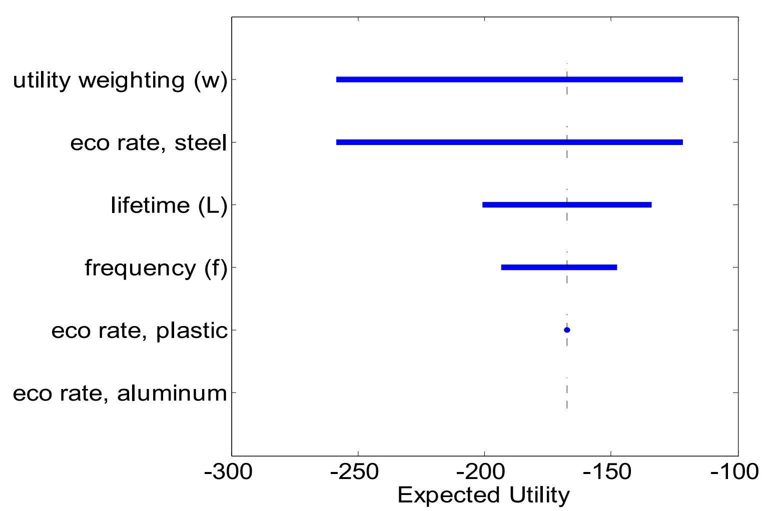


Figure 7.12. Traditional tornado diagram for one alternative

7.5 Discussion of oil filter example

In this section, PBA and decision analysis are directly compared and contrasted in four areas: veracity, acuity, complexity, and flexibility.

7.5.1 Veracity of the analysis

The example problem confirmed the assertion in Chapter 6 that a one-way sensitivity analysis can lead to the conclusion that the decision is insensitive to the uncertainty, while the PBA analysis of the same problem can indicate that the solution is very sensitive to the uncertainty. An obvious question to ask is which one gives the right result? Unfortunately, this question has no straightforward answer.

Due to repeated variables in the interval calculations, PBA gives bounds that may be overly conservative (too broad). On the other hand, one-way sensitivity analysis ignores dependencies and higher order interactions and can lead to results that are non-rigorous, i.e., that are inconsistent with the truth. In Section 6.4.1, the selection problem was compared to hypothesis testing. This comparison is now repeated for the specific example of the oil filter selection.

Consider the null hypothesis that either the PEC or the SEC filter is the best

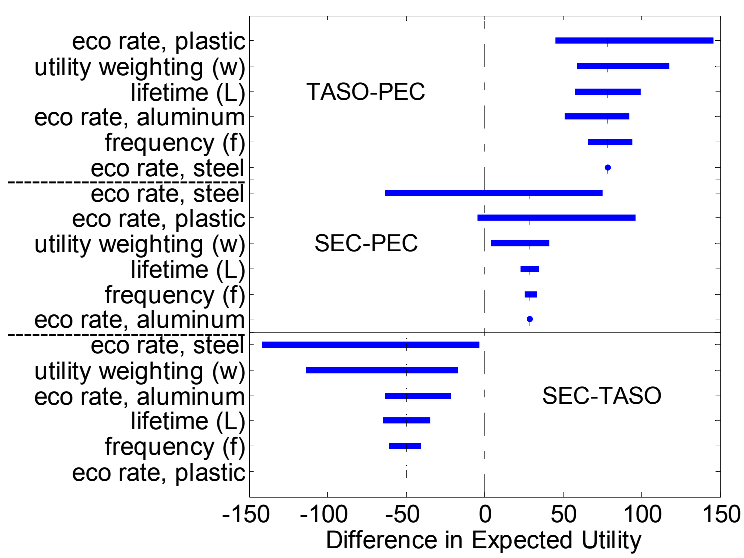


Figure 7.13. Tornado plot comparing multiple alternatives

choice. The alternative hypothesis is that the TASO filter is the best. A sensitivity analysis may underestimate the true uncertainty and indicate that there is enough evidence to reject the null hypothesis in favor of the alternative when there really is not sufficient evidence to do so. In this situation, the null hypothesis could be rejected when it is true, a Type I error. Conversely, PBA may overestimate the uncertainty and lead to the failure to reject the null hypothesis when it is false, a Type II error.

In Chapter 6, the question was posed: which is preferable, a Type I or Type II error? The oil filter example provides a context in which to why there is no general answer; the analyst must assess the situation and make his or her own choice. In this example, DASA concludes there is no sensitivity and PBA concludes there is sensitivity. Returning to Figure 6.9 and Table 6.1 on page 189, this scenario can correspond to either c_2 or c_3. If it is c_2, then PBA is making a Type II error. If it is c_3, then DASA is making a Type I error.

Considering scenario c_2, the result is that the DM concludes that no rigorous decision can be made; either more information must be collected or a policy of arbitrary choice must be adopted, even though in truth the decision is robust. This is a potential waste of resources or an unnecessary resort to arbitrary choice. Considering scenario c_3, the DM thinks the decision is robust but it in fact is not, possibly leading to a very bad decision. The DM does not know the true scenario, so he or she is left deciding which type of error is more acceptable for this particular problem. If it is inexpensive to collect additional information in comparison to the potential cost of a bad decision, then that is a reasonable route. Such cost-benefit tradeoffs in information collection are the subject of Chapter 9.

7.5.2 Acuity of analysis

One goal of sensitivity analysis is often to determine what additional information could best improve the decision. To this end, the breakout of uncertainty and sensitivity into individual quantities in one-way sensitivity analysis is an advantage. By considering each parameter independently, the decision maker gains insight into the sensitivity of the decision to each parameter.

PBA considers all uncertainties simultaneously, accounting for all interactions and dependencies, but it does not identify the individually important sources of the sensitivity. If the PBA analysis determines that the decision is not sensitive to the overall uncertainty, this is not a problem. However, in a case like Figure 7.11 in which there is indeterminacy, a DM would benefit from guidance into resolution of the indeterminacy. DASA provides this type of insight.

For example, based on the sensitivity analysis in Figure 7.13, there seems to be no need to increase knowledge about vehicle lifetime. On the other hand, the difference between SEC and TASO filter is very sensitive to the ecorate of steel, though not enough (as a one-way effect) to change the decision. The sensitivity analysis suggests that any additional information collection focus on characterizing the environmental effects. The basic PBA analysis does not provide this insight.

As discussed in Section 6.4.3, Ferson and coauthors (Ferson, et al. 2004a) have suggested using a meta-sensitivity analysis with PBA. Traditional sensitivity starts with the base values and systematically varies one parameter at a time to its extremes. Since PBA can capture all of the uncertainty at once, the opposite approach can be taken. The "base" case becomes the results with all of the uncertainty considered, and then each uncertain quantity is "pinched" down to a zero-variance interval, a precise probability, or even a point value and the reduction of uncertainty in the result is observed. However, it is not clear how to best pinch a p-box or how to measure and interpret the resultant decrease in uncertainty in the outputs. Consequently, the ability of PBA to provide guidance in questions of sensitivity of analysis for information prioritization is limited.

7.5.3 Complexity of analysis

A one-way sensitivity analysis is computationally inexpensive. In addition to the solution with the base values, each uncertain quantity requires just two calculations—one for the upper bound and one for the lower bound. Each of these calculations may involve one Monte Carlo loop to calculate expected values, although in many cases this is unnecessary. Either way, the computational complexity is generally less than with PBA.

The advantage moves to PBA as higher order (two-way, three-way, etc.) sensitivity analyses are performed, especially when nested Monte Carlo loops are used. PBA computations using dependency bounds convolutions (Williamson and Downs 1990) are generally much faster than traditional sensitivity analysis (Helton and Davis 2000, Ferson, et al. 2004a). However, as noted in Section 4.5.2, the method of dependency bounds convolutions is not applicable to black-box analysis models. Consequently, they cannot be used to analyze models such a differential equations, simulations, and finite element analysis. Current research establishes methods for propagating p-boxes through "black box" models, or models with unknown or complicated structure (Bruns, et al. 2006), and a comparison of these methods of PBA with DASA is an area of future work.

7.5.4 Flexibility of the analysis

Another advantage of PBA is its inherent flexibility. PBA's flexibility in terms of assumptions of independence or unknown dependence has already been discussed within the context of the EBDM example. Recent algorithms also handle the pairwise dependencies of maximal or minimal correlation, correlation, linear relationship and correlation within a specified interval, and signed (positive or negative) correlation (Ferson, et al. 2004a).

In this chapter, the flexibility with regard to imprecisely known distribution parameters was demonstrated, but PBA can also handle cases of unknown distribution type [see Section 6.5.1 or (Ferson and Hajagos 2004)]. This would be useful in the filter selection if, for example, the decision maker had an estimate of the mean and variance of filter change frequencies, but no theoretical or empirical evidence about the distribution family.

DASA ignores dependencies and higher order interactions, and it requires a known distribution type. Consequently, the types of problems that can be accurately explored with sensitivity analysis are more limited than PBA.

7.6 Conclusions and summary

The example decision of oil filter selection in this chapter has multiple objectives and multiple sources of different types of uncertainty. The selection problem is approached with two methods: probability bounds analysis (PBA) and traditional decision analysis with sensitivity analysis (DASA). The applicability of PBA to the example problem is illustrated and discussed in detail. This detail allows this chapter to serve as a demonstration of the process of applying PBA to a problem, as well as serving as a demonstration of the value of PBA.

Sensitivity analysis can identify important sources of uncertainty, but it can also lead to an incorrect selection because it neglects dependencies and interactions. PBA can compute with unknown distributions types, unknown dependencies, and uncertain parameters. It also provides a rigorous and global sensitivity analysis. However, PBA may yield overly conservative results (bounds that are bigger than necessary), and it is computationally more complex than simple one-way sensitivity analysis. PBA also provides little insight into information prioritization when using current methods of meta-analysis.

In short, both PBA and DASA are useful in engineering. Importantly, this conclusion identifies PBA as a useful option in engineering problems, an option that while having its own limitations, clearly reveals more information in some scenarios than traditional decision analysis. As such, these results support the conclusion that imprecision is an important characteristic of uncertainty in engineering design and that there is value in using probabilities that are most generally subjective and imprecise. In the following two chapters, attention is turned to the secondary questions of decision-making in the presence of imprecision and the principles for managing information collection.

CHAPTER 8:

DECISION MAKING IN THE PRESENCE OF IMPRECISION

According to the paradigm of decision-based design, decisions play a large role in the success of the design process. Any uncertainty model used in engineering design must therefore support decision making

When precise probabilities are used, utility theory provides a clear, normative theory for decision making. A decision-maker (DM) should choose the alternative with the highest expected value. However, when imprecise probabilities are used, expected utilities become intervals, as described in Section 4.6.1. Overlapping intervals can lead to *indeterminacy,* hence the importance of motivating question 4 addressed in this chapter:

How should a designer make a decision in the case of seemingly indeterminate information?

When there is indeterminacy, one cannot determine which decision alternative is most preferred using traditional utility theory arguments and must move to more advanced policies or resort to arbitrary choice. In this chapter, three approaches are discussed. The first involves using more of the available information to guide a choice, such as the relationship between the uncertainties of different alternatives. The second approach involves acquiring information until indeterminacy is eliminated. The final approach is to apply policies of arbitrary choice. The bulk of this chapter focuses on the first procedure: making the most of available information.

When dealing with imprecise characterizations of design alternatives, it can be useful to adopt a more set-based approach to design. In a traditional design approach, emphasis is usually placed on selecting the most preferred alternative. In a set-based approach, the emphasis moves to eliminating the least preferred alternatives. In Section 8.1, the set-based view of design is introduced. The existence of indeterminacy is revisited in Section 8.2. Policies for eliminating decision alternatives are presented in Section 8.3. The resolution of remaining imprecision is discussed in Section 8.4. The example design of a gearbox is described in Section 8.5. The various elimination policies are demonstrated using the example in Section 8.6. The process of sequentially reducing the design space is summarized in Section 8.7. The policies are discussed and future work presented in Sections 8.8 and 8.9. Most of this chapter was previously included in a conference paper (Rekuc, et al. 2006).

8.1 A set-based view of design

Design is a process of converting information about customer interests and requirements into a specification of a product. This process is complex because it involves searching through a very large, unstructured space of solutions (Tong and Sriram 1992) based on vague and uncertain knowledge about possible solution alternatives (Gupta and Xu 2002), their physical behavior (Aughenbaugh and Paredis 2004), their cost (Garvey 1999), and the decision-maker's preferences (Kirkwood and Sarin 1985, Otto and Antonsson 1992, Carnahan, et al. 1994, Seidenfeld, et al. 1995). The complexity of the design problem, including the presence of uncertainty, makes it impossible to arrive at a final design in one step, as described in Section 1.1.2.

Researchers have recognized the limitations of sequential design processes and have proposed modifications in which the uncertainty about future design decisions is considered. For instance, Chen, Allen, and coauthors (1996) have introduced an approach based on robust design that seeks decision alternatives that are robust to future decisions. The idea is that since DMs lack knowledge about the outcomes of future design decisions, they should make their current decision in a way that yields a *satisficing* solution (Simon 1982)—a solution that is in some sense good enough regardless of future design decisions.

Robust design methods trade off optimal performance for consistent performance. This is a reasonable approach if the price one pays for robustness is relatively small—that is, if little performance is sacrificed for robustness. Unfortunately, robust design methods currently do not explore how large that price is. In this chapter, an approach is demonstrated that helps the DM move toward the most preferred solution by actively managing the design space, rather than compromising high performance for robustness. This approach is inspired by set-based concurrent engineering.

The perspective taken in this chapter is that ideally the goal of the design process is to systematically eliminate inferior design alternatives from the set under consideration until only the most preferred alternative (or set of equally preferred alternatives) remains. This approach is derived from the paradigm of Set-Based Concurrent Engineering (SBCE) (Sobek, et al. 1999), a management approach used at Toyota. The guiding principle of SBCE is to begin the design process by selecting a broad set of solutions and gradually narrowing the set by eliminating weaker solutions as more information becomes available until converging on a final solution.

In traditional design practice, the emphasis is on selecting a single good design; engineers quickly converge on a single design and then iteratively modify that solution until it meets the design requirements. In SBCE, Toyota encourages its engineers to pursue multiple feasible design alternatives simultaneously. The consideration of multiple designs incurs more costs early in the design process than selecting a single robust design. However, these increased costs can be offset by two factors. First, the resulting design in SBCE is often much closer to optimal (has a much better performance) than the final designs in traditional methods. Second, SBCE avoids costly design tweaking and redesigns late in the development cycle. In some combination, these effects have enabled Toyota to use SBCE quite successfully (Parunak, et al. 1997, Sobek 2004).

As implemented at Toyota, SBCE places a large responsibility on chief engineers to guide the process effectively, relying heavily on their implicit knowledge and expertise. Sobek, Ward, and Liker (1999) provide three broad principles for managing SBCE. One of these principles involves narrowing sets gradually while increasing detail, and a formal method has been introduced that uses predicate logic to eliminate infeasible designs (Finch 1997). However, little attention has been given to methods that guide elimination based on preferences—that is, methods that eliminate less-desirable, yet feasible, designs from the set under consideration. If the benefits seen at Toyota are to be generalized to other applications, a formal method of set-based design must be developed.

For a set-based approach to be efficient, the DM must be efficient at eliminating inferior solutions from the set under consideration; a DM should eliminate a solution as soon as he or she is confident that it cannot be the most preferred. If the DM does not eliminate such solutions, then he or she will continue to develop them in more detail, thereby incurring unnecessary costs. Since these elimination decisions must be made without complete knowledge about the solutions, traditional comparisons are inappropriate; different methods are needed. The remainder of this chapter discusses elimination decision policies and demonstrates an elimination-oriented, set-based design approach using the design of a gearbox as an example.

8.2 Indeterminacy in Decision Making

In this section, the scenario of overlapping intervals of expected utility, mentioned throughout the dissertation, is revisited. In general, there are three possible scenarios of preference between alternatives A and B. Either A is preferred to B, B is preferred to A, or the DM is indifferent between A and B. When utilities are used to reflect preference, these relationships can be determined by the inequality or equality of the expected utilities (von Neumann and Morgenstern 1944). However, when imprecision exists, the expected utilities become intervals (since the probabilities are not uniquely determined, neither is the mathematical expectation), and such comparisons become more complicated.

For example, consider the intervals of expected utility for two alternatives (A and B) shown in Figure 8.1(a), which is a repeat of Figure 4.8(a). In this example, the intervals overlap and the relationship between the alternatives cannot be stated in terms of standard mathematical comparisons such as $A > B$, $A < B$, or $A = B$. Since the true expected utility of B can lie anywhere in the given interval, the point labeled b_1 is possible. Similarly, both a_1 and a_2 are possible true values for the expected utility of A. Notice that a_1 is greater than b_1, but a_2 is less than b_1. Consequently, the available evidence is *indeterminate*; the DM cannot determine which alternative is the most preferred, nor can the DM determine that he or she is definitely indifferent. In order to make elimination decisions in the presence of imprecision, different methods are needed.

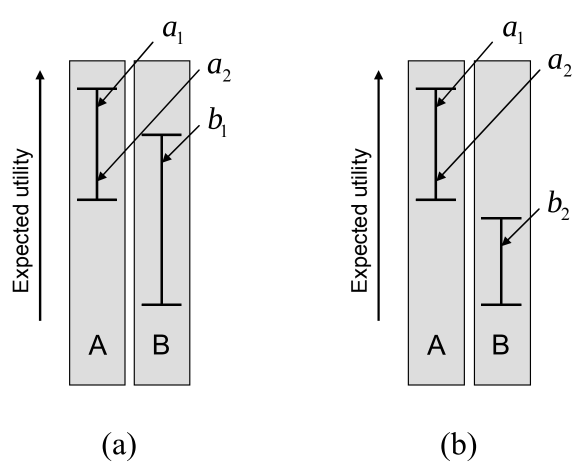


Figure 8.1. Intervals of expected utility

8.3 Elimination decisions with imprecise information

As demonstrated at the close of the previous section, standard numerical comparisons are insufficient for elimination under imprecision. Instead, a DM must turn to interval methods such as interval dominance, maximality, or E-admissibility.

8.3.1 Interval dominance

An example of overlapping intervals was shown in Figure 4.8(a). Obviously, two intervals will not always overlap. In this case, shown in Figure 4.8(b), it does not matter where in the given interval the true expected utility of A falls—it will always be greater than any value in the interval for expected utility of B. This illustrates a situation referred to as *interval dominance* [for a brief synopsis, see (Zaffalon, et al. 2003)]. Interval dominance can be defined as follows. Consider two intervals $A=[\underline{A},\overline{A}]$ and $B=[\underline{B},\overline{B}]$. Interval A dominates interval B if and only if $\underline{A}\geq\overline{B}$, that is, if the lower bound on A is at least as high as the upper bound on B. By extension, this condition means that there is no point in interval A that is lower than any point in interval B.

Interval dominance is obvious when there are only two alternatives, but it is more subtle when there are more alternatives, such as shown in Figure 8.2. At first glance, it may appear that no elimination is possible because there is significant overlap between intervals. However, by comparing all pairs of alternatives, one discovers that alternative D is dominated by alternative A, and hence can be eliminated. By using as a reference for comparison the alternative with the maximum lower-bound and then comparing this

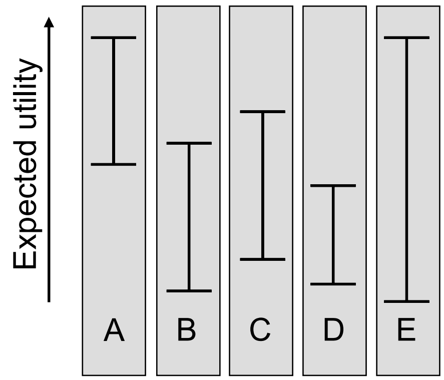


Figure 8.2. Many overlapping intervals

to the upper-bounds of all other alternatives[16], the complexity of the calculation is reduced from $O(n^2)$ to $O(n)$ for a problem with n alternatives. The result of applying this criterion is a set of alternatives whose intervals of expected utility all share some region of overlap.

Elimination using interval dominance s is relatively easy to compute, but it may result in a large set of design alternatives. Although this is to be expected, especially during the early phases of design, it is important for the success of this approach that as many designs as justified by the available knowledge and information are eliminated;

---

[16] Essentially, the problem is reduced to two loops over all the alternatives, avoiding combinatorial evaluation. A simple pseudo-code example illustrates the method:

Define maxmin = some large negative number
For each alternative i
  If maxmin < minimum value in interval of expected utility for alternative i ($minEU_i$),
  Then set maxmin = $minEU_i$.
End for
For each alternative i
  If maxmin > maximum value in interval of expected utility for alternative i ($maxEU_i$),
  Then eliminate alternative i from consideration.
End for

inefficiencies should be avoided. In the next section, two approaches that account for uncertainty shared across alternatives, the criteria of maximality (Walley 1991) and E-Admissibility (Levi 1974), are considered.

8.3.2 Accounting for shared uncertainty

In design, there are often uncertain quantities that influence the performance of all decision alternatives in a similar fashion. Such quantities are defined as *shared uncertainties*. For an uncertain quantity to be considered shared between two design alternatives, two conditions must hold:

1. The realization of that uncertain quantity must be independent of the action chosen by the DM (i.e. act-state independence must hold)
2. The consequences of both alternative actions must be a function of the uncertain quantity

For example, ambient temperature is independent of the alternative chosen—all potential final designs will have to operate over the same, but unknown, range of temperatures. Hence, the uncertainty is said to be shared. As an example of uncertainty that is not shared, consider the sequential decisions of designing first the engine of a car and then the drive shaft. When designing the engine, the exact design of the drive shaft is unknown. However, this uncertainty is not shared by all engine alternatives, because the final design of the drive shaft will depend on the chosen engine design; for example, the drive shaft must meet difference performances requirements depending on the power of the engine.

Since temperature is a shared uncertainty, the performance of alternatives should be compared assuming they are operating at the same temperature. A similar argument favors paired statistical testing over pooled statistical testing to remove shared systematic errors (Devore 1995). The motivation is illustrated in the following example.

Consider two cars A and B, whose performance each depends strongly on the uncertain ambient temperature T, such as shown in the top left of Figure 8.3. Note that for all values of the uncertain parameter, the utility of A is greater than the utility of B. Clearly then A is the superior alternative. However, if only the intervals of performance were compared without regard to shared uncertainty, such as shown in the top right of Figure 8.3, this superiority would not be detected.

The concept of shared uncertainty is similar to common random numbers (CRNs) in discrete-event simulations (Law and Kelton 2000). The goal of a simulation is usually to compare two scenarios or alternative designs by examining the difference in output for different combinations of control parameters. If different random numbers are used in the simulations for the different alternatives, additional noise is introduced into the model. CRNs are used to induce correlation between scenarios, thereby reducing the variances of the results.

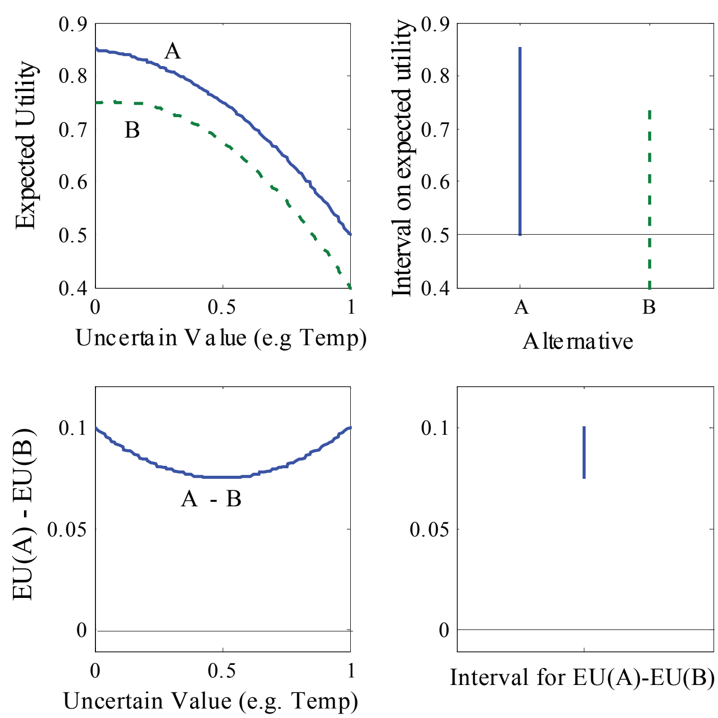


Figure 8.3. Comparing two alternatives with and without shared uncertainty.

In engineering design, shared uncertainty is an inherent characteristic of the problem. Therefore, a DM does not have to add the commonality; he or she merely needs to recognize it and to take advantage of that additional property when it exists. One approach that considers shared uncertainty is the maximality criterion.

8.3.2.1 Maximality criterion

Since the intervals in the top right quadrant of Figure 8.3 overlap significantly, neither A nor B is eliminated according to interval dominance, even though it is clear from the curves in the top left quadrant that B should be eliminated. If the difference in performance across the uncertain parameter is considered, the elimination can be made, as shown in the lower left quadrant of Figure 8.3. Note that for any and all values of the shared uncertain variable, the difference between alternative A and alternative B is positive. In other words, A is always better than B. Consequently, B can be eliminated; it cannot be the best alternative because it is inferior to A at all temperatures.

This type of comparison is formalized as the *maximality* (Walley 1991) criterion for elimination. Maximality is defined as follows. First, a set of decision alternatives that are available for consideration is defined and denoted D. Next, shared uncertainties (denoted $z_s \in Z_s$) are distinguished from uncertainties that are specific to each alternative (denoted $z_i \in Z_i$ for alternative $A_i \in D$). Recalling that the DM seeks to maximize expected utility, the elimination rule of maximality is defined as follows:

> A decision alternative $A_i \in D$ is dominated according to maximality, and hence the corresponding set of design alternatives can be eliminated, if for some $A_j \in D$, and $i \neq j$:

$$\left(\max_{\substack{z_s \in Z_s \\ z_i \in Z_i \\ z_j \in Z_j}} [EU(A_j, z_j, z_s) - EU(A_i, z_i, z_s)]\right) < 0$$

Maximality is a stricter criterion than interval dominance, meaning that in general it leads to the elimination of more alternatives. Maximality eliminates alternatives that are dominated at all values of the uncertain parameter by any individual other alternative. In general, this requires the maximality condition to be checked for all pairs of decision alternatives. For example, consider the five decision alternatives whose expected utility is expressed as a function of a single shared imprecise parameter z_s (for example, ambient air temperature, $z_s = T$) in Figure 8.4. If one were to use only A as a *reference design* (meaning comparing the other designs only to A), then only B could be eliminated because curves C, D, and E are higher than A for some values of z_s, but B is always lower. However, if C is used as the reference design, then D is also eliminated. Clearly to complete the elimination, both A and C must be used as reference designs in this case. In general, a DM must compare all combinations.

The difference in expected utility is often monotonic with respect to the uncertain variables. In this case, the maximum difference occurs at the boundary of the uncertainty region, making it easy to compute for a given pair of alternatives A_i and A_j. If the

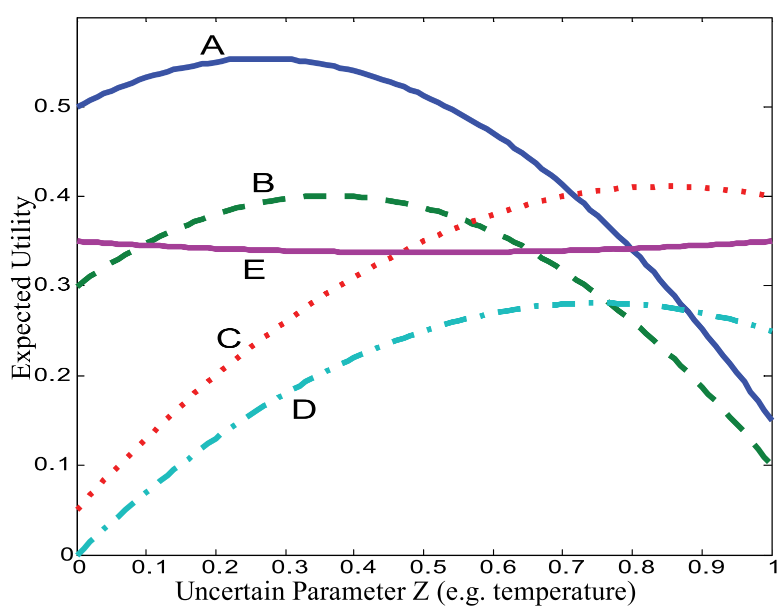


Figure 8.4. Performance of 5 alternatives influenced by a single uncertain parameter (e.g. temperature).

difference between the two alternatives is not monotonic then a complete optimization is necessary. This obviously increases the cost of applying maximality beyond that of interval dominance. The tradeoff is in the more complete elimination of truly dominated alternatives.

8.3.2.2 E-admissibility criterion.

A stricter criterion than maximality is E-admissibility (Levi 1974). According to E-admissibility, a solution is eliminated if at every value of the uncertain parameter there is at least one alternative with a higher expected utility. This is more easily understood by considering the converse—E-admissibility only accepts alternatives for which there is at least one value of the uncertain quantity at which no other alternative has a greater expected utility. In other words, an alternative is accepted only if it is optimal at some value of the uncertain quantity. Applying E-admissibility in general requires solving a mathematical programming problem, or at least proving that a feasible solution exists (Kyburg and Pittarelli 1996), making it at least as expensive as applying maximality.

For an example of applying E-admissibility, consider the alternatives in Figure 8.4 once again. Alternatives eliminated using maximality are necessarily eliminated using E-admissibility since if at all temperatures there is a single alternative with higher expected utility, then there can be no value of the uncertain quantity for which there is no alternative with a higher expected utility. Consequently, alternatives B and D are eliminated using E-admissibility.

Notice that alternative A performs best in low temperatures, C performs best in high temperatures, while E performs consistently throughout the entire temperature range. Alternative E can be considered a robust solution, where a robust solution is one that can be exposed to variations in the environment (and other factors) without suffering unacceptable performance degradation (Allen, et al. 2006). Nevertheless, E will be eliminated based on the E-admissibility criterion because E is dominated by the set

$\{A, C\}$, since either A or C (or both) is greater than E at every temperature; car A dominates E at low temperatures while C dominates E at high temperatures. There is no value of the ambient temperature at which there is no alternative with a higher expected utility than that of E. The potential implications of eliminating the robust solution E are described in the following section.

8.4 Resolving remaining imprecision

Although a DM can maintain a set of designs in the early stages of design, he or she must eventually select a particular alternative to finalize the design. After applying elimination criteria, multiple alternatives usually will remain due to imprecision. In order to make a final decision, a DM has two choices—the DM can collect additional information, thereby reducing imprecision until only one alternative remains in the non-dominated set, or the DM can select one alternative arbitrarily. Traditional design approaches would require arbitrary elimination of non-dominated alternatives, while a set-based design approach allows a DM to delay elimination of alternatives until additional information is available.

When delaying decisions, the DM should carefully consider the tradeoff between the value of obtaining more information and the cost of doing so by applying information economics [see Chapter 9 and (Ling, et al. 2006)]. Although the cost of additional investigation is often worth the improved ability to make a more informed decision, the DM will reach a point at which the cost of gathering additional information outweighs the expected benefits. At that point, the DM should resort to the other option: arbitrary choice.

If a DM is unable to resolve the imprecision before needing to choose a single alternative from the set, he or she may need to make an arbitrary choice—a choice that is not uniquely determined by the DM's preferences, beliefs, and values (Walley 1991). Recall that the presence of indeterminacy implies that the available information does not uniquely identify a most preferred alternative. Consequently, any arbitrary choice from among indeterminate alternatives can be defended as rational in the context of a single decision.

Arbitrary in this sense does not necessarily imply *without guidance* or *random*. Several policies are of arbitrary choice have been proposed, including Γ-maximin (Berger 1985) and the Hurwicz-criterion (Arrow and Hurwicz 1972). A Γ-maximin policy says that given indeterminacy in a maximization problem, a DM should select the alternative with the highest lower-bound. This is a conservative policy in that it seeks to mitigate the worst-case. Robust design strategies that choose solutions that are insensitive to imprecision are also applicable at this stage. If the remaining uncertainty is extreme, it may be valuable to consider an alternative approach such as information gap theory (Ben-Haim 2001).

Returning to the vehicle design example in Figure 8.4, assume that the DM is unable to resolve the imprecise ambient temperature, yet has to choose between the alternatives (A and C) that remain after applying the E-admissibility criterion. The Γ-maximin policy would choose alternative A, because it has the highest lower-bound over the range of the uncertain parameter. However, notice that while A performs well in low temperatures, it performs quite badly at high temperatures. Notice also that alternative E performs moderately well at all temperatures—that is, alternative E is robust to temperature. However, using the E-admissibility criterion, this alternative was eliminated.

The elimination of alternative E is not a problem if the DM is able to resolve the imprecision before needing to choose a final design. Once the imprecision is eliminated in this example, the DM knows at which temperature the car is required to perform and can select the car that is most preferred at that temperature—which will always be either car A or C, the two alternatives that remain after applying the E-admissibility criterion. Therefore, if a DM knows that imprecision can be eliminated before making the final decision [an example of such imprecision is that from future decisions], then E-admissibility is an appropriate criterion.

However, information economic considerations will usually lead a DM to stop collecting information before removing all imprecision. In other cases, such elimination of imprecision is impossible. For example, there is in general no one temperature at which a care must operate, but rather a produced car could be subject to the entire range of temperatures during operation. The practical unlikelihood of removing all imprecision leads us to recommend the maximality elimination criterion for most design applications.

8.5 The design example

In this section, gearbox design problem is used to demonstrate the different elimination criteria. The gearbox is intended for use in the drivetrain of an SAE Mini-Baja competition off-road vehicle. The basic configuration of the gearbox is shown in Figure 8.5, and the objective of the design problem is to determine the geometries of the three gears such that the expected utility of the design is maximized.

A summary of the problem formulation is presented in Figure 8.6 on page 239. Utility is formulated as the dollar earnings from constructing and using the gearbox in Georgia Tech's Mini-Baja vehicle for a long-distance race. There are five design variables and ten shared uncertain parameters, with uncertainty modeled as p-boxes, intervals, and precise probability distributions.

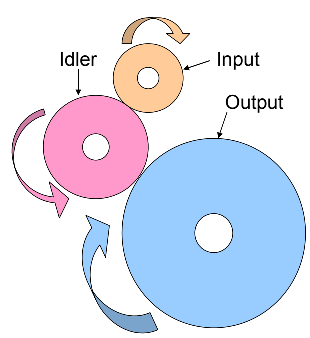


Figure 8.5. Gearbox configuration schematic

As noted in Chapter 4, probability bounds analysis (PBA) is a refinement of imprecise probabilities in which uncertain information is represented as a p-box. While less expressive than imprecise probabilities, the p-box representation simplifies computation. In particular, DBC allow for the computation of rigorous, "best-possible" bounds on functionally determined distributions (Williamson and Downs 1990, Berleant and Goodman-Strauss 1998). Unfortunately, DBC is often inappropriate for realistic engineering problems because it cannot be applied directly towards a black box analysis model.

An alternative to these methods is to use an entirely stochastic approach. For precise probabilistic problems, Monte Carlo sampling can be used to approximate the uncertainty in the output. For imprecise probabilistic problems, second-order (2D), also known as two-dimensional or double-loop, Monte Carlo sampling can be used to approximate the imprecise uncertainty in the output (Hoffman and Hammonds 1994). However, for the high dimensional problems typical in engineering design, computational expense could still be prohibitive.

For the gearbox example problem, an alternative approach is used in which the outer loop of a 2D Monte Carlo simulation is replaced with an optimization algorithm [see Bruns and co-authors for more information (Bruns 2006, Bruns and Paredis 2006, Bruns, et al. 2006)]. The inner loop remains a Monte Carlo sampling from parameterized distributions, but instead of determining the parameters of these distributions by an outer loop of Monte Carlo sampling, optimization is used to find the set of distribution parameters that will give us the largest (and the smallest) expected values. The results of these calculations are then used to make the comparisons in the elimination criteria.

Maximize

Expected Utility:

$EU = E[U_t] * P\{\text{complete}\} - E[U_C]$

where

- $U_t = (\text{Prize Money}) * \left(1 - \frac{1}{1+e^{16-4t}}\right)$, with the relationship determined by fitting a sigmoid function to past race finish times t.
- $P\{\text{complete}\}$ is the probability that the gearbox completes the race, i.e. the reliability
- U_C =the cost of constructing the gearbox

Select

Gear Ratio $N_g = [0.5, 5]$ (torque ratio)

Input Gear Diameter $d_{in} = [1.5, 15]cm$

Idler Gear Diameter $d_{id} = [1.5, 15]cm$

Gear Width $w = [1.00, 8.75]cm$

Gear Module $M = [1.27, 8.75]mm / tooth$

Where

Performance depends on 10 uncertain system parameters shared across all alternatives:

Total Mass (kg), $M \sim Normal([200, 215], [18, 20])$

External Drag Coefficient (N/(m/s)2), $C_{D,e} = [0.27, 0.28]$

Internal Drag Coefficient (N/rpm), $C_{D,i} = [0.0, 0.0075]$

Course Roughness Coefficient, $K_C \sim Normal(3, 0.5)$

Bending Strength Factor, $J = [0.38, 0.4]$

Gear Quality, $Q_v \sim Normal([8.25, 8.75], 1)$

Cost Error ($), $Cost_{err} = [-5, 5]$

Uncorrected Bending Strength (N/m2),

S'_{fb} ~ Normal([197,203]x106, [30,35]x106)

Uncorrected Contact Strength (N/m2),

S'_{fc} ~ Normal([197,203]x106,[30,35]x106)

Application Factor, $K_a = [1.68, 1.70]$

Figure 8.6. Formulation of Mini-Baja gearbox problem.

This approach, while computationally more efficient, assumes independence between uncertain variables. While not ideal, this assumption is reasonable for large classes of engineering models. Nevertheless, this assumption removes one of the advantages of p-boxes, that being the flexibility of one algorithm (DBC, see Section 4.5.1) to compute with unknown dependencies. Another limitation of this approach is the presence of local minima in typical engineering problems. For the gearbox example problem, this difficulty was overcome by using several starting points for the optimization. However, this approach would not be appropriate for more complicated problems. The need to start the optimizer with several initial points and the inability of the optimized computational approach to consider cases of non-independent inputs indicate directions for future work.

8.6 Demonstration of existing elimination criteria

The first part of the example demonstrates the reduction of the design space for the first design variable—the gear ratio. This design problem is slightly different from the simple examples mentioned earlier in the chapter because it deals with a continuous variable. For a continuous design variable, it is ranges of values that are eliminated, rather than discrete alternatives. The initial problem statement specifies the design space of gear ratios in the interval [0.5, 5.0]. In the first step of the sequential decision process, the DM seeks to reduce this interval as much as possible while retaining in the range the most preferred—though currently unidentifiable—solution.

The application of *interval dominance* by the DM is considered first. Figure 8.7 contains a plot of expected utility versus gear ratio. The two curves represent the upper and lower bounds on expected utility at a given gear ratio. In the plot, the highest point on the lower-bound, or the Γ-maximin solution, occurs at a ratio of about 1.5. The DM draws a horizontal line at the lower expected utility at this gear ratio. By the condition of interval dominance, any gear ratio with an upper-bound on expected utility that is below this line should be eliminated. For example, two expected utility intervals are indicated in Figure 8.7. The leftmost interval is located at the Γ-maximin solution. The DM compares all other decision alternatives to this interval. The Γ-maximin solution clearly dominates any of the expected utility intervals in the shaded regions. Therefore, the DM can eliminate gear ratios in both shaded regions from the design space.

By taking into account that the uncertain quantities described in Figure 8.6 are shared between different design possibilities, further eliminations in the design space can be made using the maximality criterion. In theory, the DM would need to make pairwise comparisons between all alternatives to eliminate all that are dominated under maximality. Of course, this is impossible for design problems with continuous design variables. In practice, a DM must therefore perform comparisons between a well-chosen discrete set of design alternatives the effectively sample the entire design space.

In this example, the DM computes the bounds on the expected difference in utility between each gear ratio and the Γ-maximin gear ratio of 1.5. Recall that the maximality

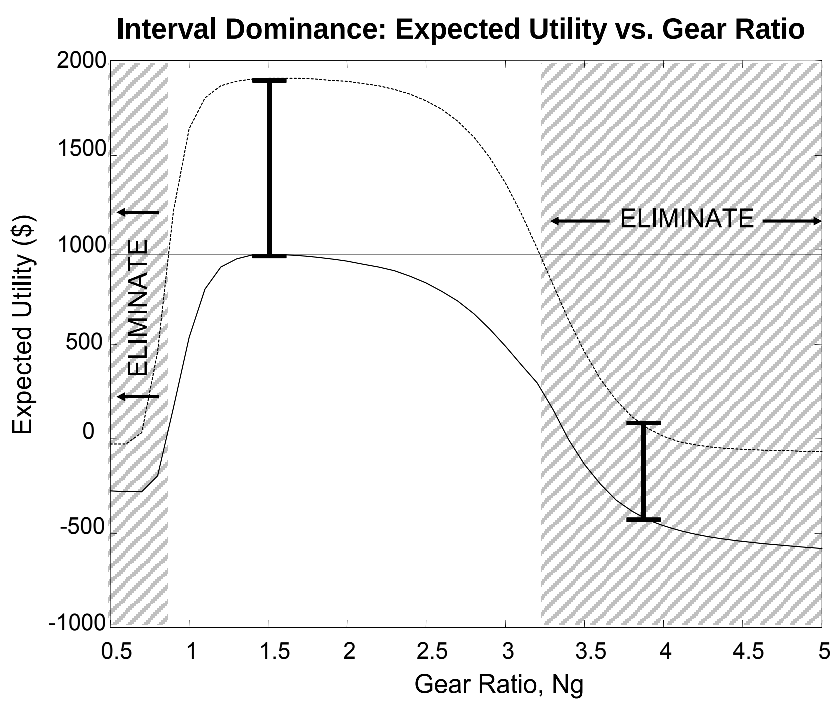


Figure 8.7. Elimination using interval dominance

elimination criterion specifies that the DM should eliminate any alternative (in this case, gear ratio) with an upper bound on expected difference less than zero. Figure 8.8 contains a demonstration of *maximality* elimination over a continuous variable. The two curves in the figure represent upper and lower expected *differences in utility* between the current alternative and the reference material, in this case defined to be a gear ratio of 1.5. The DM draws a horizontal line at an expected difference in utility of zero. The shaded regions correspond to gear ratios that are always dominated by designs with the reference gear ratio of 1.5. Therefore, the DM can eliminate all decision alternatives that fall in the shaded regions in Figure 8.8.

The calculation of a difference in expected utility requires two alternatives—the one being tested, and a reference. In order to increase the efficiency of elimination, we choose a detailed reference design (Rekuc 2005). The idea is to develop one promising alternative to a greater level of detail than the others, thereby reducing the imprecision from future decisions for that alternative. The narrower intervals of utility for this design will often enable more

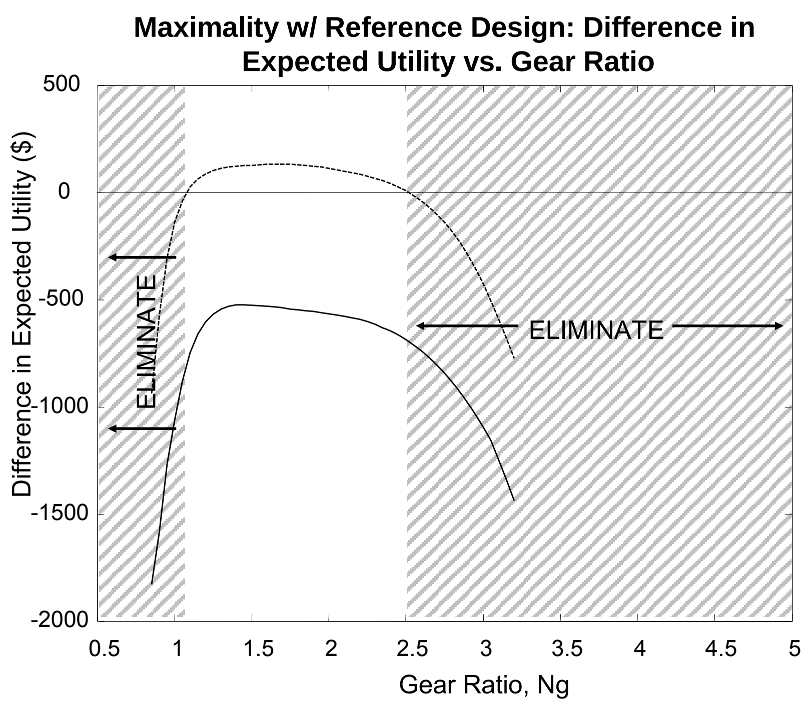


Figure 8.8. Eliminating using maximality

elimination. In this paper, the Γ-maximin solution is used as the reference design. Specifically, comparisons are made to the reference design of $N_g = 2.1$, $d_{in} = 1.5cm$, $d_{id} = 1.5cm$, $w = 1.25cm$, and $M = 6.35mm$. The mechanics of the resulting elimination decisions are the same as those described earlier in this paper.

8.7 Sequential reduction of the design space

The examination of the example problem is concluded by considering the sequential process of reducing the set of feasible designs, as sketched in Figure 8.9. A single step in this process was described in the last section, in which the design interval for the gear ratio was reduced. Next, the set of non-dominated design alternatives is reduced sequentially for each of the remaining four design variables.

The advantage of sequential elimination is that with each reduction in the uncertainty associated with a single design variable, the uncertainty in expected utility is reduced. Often, the reduction of uncertainty at one step allows for additional reduction of

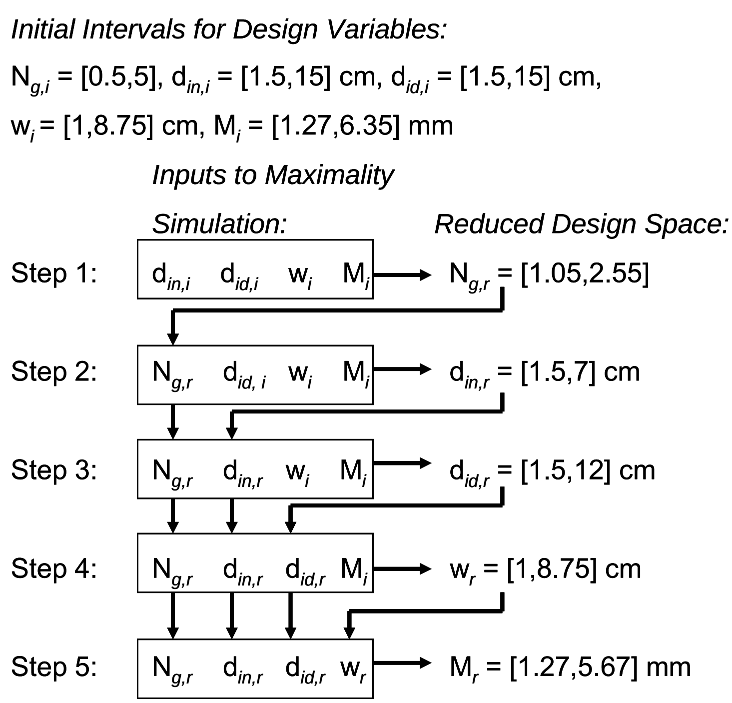


Figure 8.9. Sequential reduction process.

uncertainty to be made in the next step.

In step 1, the DM reduces the interval for the gear ratio based upon the initial design uncertainty for the other four design variables. In step 2, the DM reduces the uncertainty for input gear diameter based upon the reduced uncertainty for the gear ratio and the initial uncertainty for the other three design variables. The DM repeats this process sequentially until the design spaces for all design variables have been reduced via elimination. The process could then be repeated for further reductions. The right side of Figure 8.9 contains the intervals representative of the reduced design space for one iteration. Additional aspects of this process are addressed in the next section.

8.8 Discussion of results and identification of future work

This chapter contains a motivation for and a demonstration of a sequential, set-based design approach in which the decision maker (DM) explicitly considers imprecision. In this process, the DM incurs extra costs in the form of additional computation time and the expenditure of additional resources for developing and evaluating sets of design alternatives. In exchange for these costs, the DM receives the benefit of converging on the most preferred design alternative and avoiding costly redesign. Before adopting these methods, a DM should consider whether these benefits outweigh the costs. This question requires further research, and the answer to the question will likely depend on the development of efficient means for managing and organizing the sequence of decisions in set-based design.

Specifically, it would be valuable to develop a formal set-based design model that goes beyond the general management principles of SBCE (Sobek, et al. 1999), feasibility-based elimination (Finch 1997), and the elimination methods presented in this paper. Such a model would define the partitioning of the design problem into sets, the sequence of decisions, and definitive rules for elimination and arbitrary selection. For example, the order in which the design variables are explored in Figure 8.9 may affect the results.

It was noted earlier that one factor in choosing between the maximality and E-admissibility criteria was whether or not imprecision can be eliminated before a final decision is made. In practice, it costs resources to eliminate imprecision, and resources are always limited. At some point, the cost of additional information collection will exceed the expected benefit in increased performance of the design solution. A DM must choose when to stop expending resources to eliminate imprecision and to select a final design solution arbitrarily, taking guidance from information economics in engineering design (see Chapter 9). However, the issue remains as to how to make this final arbitrary decision in a manner that works effectively in different classes of engineering problems, if such a method even exists.

There is also substantial room for improvement in the computational method for propagating imprecise probabilities through the model. The method used in this example models all uncertain variables as independent. This may be sufficient for certain problems, but how would the computations change for dependent or correlated uncertain variables? In addition, with the current method, it is not clear how close the computed upper and lower expected utilities are to the actual expected utilities or to the so-called rigorous, "best-possible" bounds for unknown dependence between the uncertain inputs. Finally, more efficient computational methods would be necessary for uncertainty propagation in complex, computationally expensive models.

8.9 Summary

In this chapter, the issue of decision making with imprecise information was addressed. The value of adopting a set-based view of design was introduced, and several decision criteria were presented. Specific new contributions are the identification of the limitations of the E-admissibility criterion in engineering design and the subsequent recommendation of the maximality criterion for engineering design decisions. The policy of maximality was recommended for use in engineering design. As an example context, a gearbox design problem involving multiple design variables and sources of uncertainty was presented. Using this example, the ability of interval methods to guide the elimination of alternatives in set-based design was demonstrated. Finally, priorities for future work on decision making with imprecise information were identified.

CHAPTER 9:

BOUNDING THE VALUE OF FUTURE INFORMATION COLLECTION

As described in Section 1.1, engineering design is a sequential and iterative process, consisting of several phases: product planning and clarification of task, conceptual design, embodiment design, and detail design (Pahl and Beitz 1996). Decision-based design research recognizes a sequence of decisions in the design process and emphasizes the importance of these decisions to the success of the design (Thurston 1990b, Hazelrigg 1996, Marston, et al. 2000). Each decision has two phases—problem formulation and problem solution. Problem formulation consists of identifying design alternatives, states of the world, probabilities of the states, payoffs for each alternative under every state, and criteria for evaluation (Kmietowicz and Pearman 1981). Part of the process of identifying the probabilities of the states of the world is a sub-decision problem regarding how much information to collect—the subject of this chapter, which is based heavily on a published conference paper (Aughenbaugh, et al. 2005) and forthcoming journal paper (Ling, et al. 2006). Much of this work was also included in (Ling 2006).

9.1 Decision formulation and information collection

Engineers support decision making throughout the design process by expending resources to create, collect, and analyze information. Although this information may provide value to the designers by leading to a better final design, this benefit is uncertain until resources are spent and the information is actually collected and used. In this chapter, the incorporation of the management of this cost-benefit tradeoff into the design decision model is addressed.

In related research, Gupta and co-authors have demonstrated the importance of incorporating the cost (in terms of number of design alternatives considered) of decision making into the overall design decision model (Gupta and Xu 2002), but they do not provide an approach for estimating the value of information in actual design problems. Radhakrishnan and McAdams consider the cost-benefit trade-offs in selecting models of various levels of abstraction in engineering design (Radhakrishnan and McAdams 2005). They present a framework in which a designer can reason about model uncertainty, but they admit that the designer is left with little guidance in estimating the actual value of information from different models. Along similar lines, Bradley and Agogino develop a decision-analytic approach to assist designers in cost-benefit analysis of resource expenditures using precisely characterized probability distributions to guide and prioritize information collection (Bradley and Agogino 1994), but they do not explain how to estimate these distributions.

In the simulation literature, statistical output analysis is commonly performed to assess whether a sufficient number of simulation replicates have been performed to obtain statistically significant conclusions (Law and Kelton 2000). Standard practice in simulation output analysis is to specify a requirement on accuracy of predictions *a priori*. Then some initial number of simulations are run in order to get rough estimates of the standard deviation of the output quantities. This estimate of the standard deviation is used to estimate how many more simulation runs are required in order for the estimates of the means of the output quantities to converge to the required accuracy.

Because this analysis uses a predetermined measure of accuracy, several important aspects are lost. First, the desired accuracy is determined before the problem is explored, meaning that it set without the guidance of the analysis. For some problems a high accuracy is needed, while for others a low accuracy is sufficient. Second, the cost of creating and running additional or more complex simulations is not considered as part of the problem. Even if there is benefit in creating and performing more simulations, this benefit may not exceed the costs.

As in any kind of cost-benefit analysis (Layard and Glaister 1994), a common unit of measure is needed. This need can be met by using the economic value of information (Lawrence 1999). Although the economic value of information is clearly correlated with accuracy, they are not equivalent. The cake or death example of Section 1.3 illustrated that sometimes very coarse, inaccurate models can still be useful for decision making.

For example, when distinguishing between two alternatives that differ significantly in performance, a very accurate and expensive model is less valuable than a simpler model that could have made the same distinction at a much lower cost. Conversely, in high-risk design problems, an expensive model that is more accurate than typically required may lead to a better solution even when factoring in costs, since a simple model may lead to a decision with disastrous consequences.

Howard develops a theory of the value of information which takes both probabilistic and economic factors into account and he uses this theory to determine the optimal number of tests to perform to characterize a known distribution with unknown parameters (Howard 1965, Howard 1966). Matheson extends Howard's theory and uses it to determine the most economic computations and analyses to perform for a particular decision problem (Matheson 1968). Although Howard's and Matheson's works are similar in objective to this paper, their approach depends on the designer's ability to accurately assign precise probabilities to the possible states of nature (i.e. having accurate priors) before performing the analysis.

In this paper, imprecise probabilities are used to extend the applicability of information economics to cases in which probability distributions are not perfectly known to designers—specifically, to the management of statistical data collection in support of engineering design decisions. This new method is illustrated by extending the pressure vessel example of Chapter 5 .

9.2 Information economics in engineering design

In this section, the basics of information economics are introduced. Economics is the study of choice under conditions of scarcity (Lieberman and Hall 2000). Extending this definition, information economics is the study of choice in information collection and management when resources to expend on information are scarce. Because designers face a scarcity of resources, such as time and money, the principles of information economics should be applied to the information collection process in engineering design.

The area of information economics grew out of statistical decision theory in the 1950s when Marschak published a series of papers on the economics of information and organization (Marschak 1974). Recently, with the explosion of new information technologies, information economics has regained attention within the broader context of information management. Current areas of research focus on corporate finance and industry policy, such as intellectual property rights, industry regulation, and fostering innovation (Rubin 1983, Strassmann 1999), or on the infusion of information technology into a corporation (Strassmann 2004). Within engineering, the focus of information management has been primarily on data exchange, interoperability, and visualization to support collaborative design. For an overview of these areas, refer to the following review articles (Ciocoiu, et al. 2001, Jayaram, et al. 2001, Rangan and Chadha 2001, Szykman, et al. 2001, Urban, et al. 2001).

In a more general sense, information economics presents principles by which the cost-benefit tradeoffs of information collection can be managed in engineering design. Many of these principles have been developed and employed previously in standard micro-economics and the theory of the economic value of information, pioneered by Marschak (Marschak 1974) and summarized by Lawrence (Lawrence 1999). A substantial difference between engineering design applications and those of Marschak and Lawrence is the availability of perfectly known probability distributions—knowledge that Marschak and Lawrence assume to be available, but engineers often lack in practice. The goal of this chapter is to apply information economics directly to the management of information and uncertainty in engineering design.

9.3 Example problem

Throughout the remainder of this paper, the application of information economics is discussed in the context of an example of a pressure vessel design. This example has been used previously to demonstrate the value of using imprecise probabilities in engineering design decisions in Chapter 5. This experiment is now extended to explore the decision of how much information to collect in order to support design decisions.

In the modified example problem, a pressure vessel is designed to meet certain requirements while maximizing payoff. The complication is that the pressure vessel is to be built using a material with unknown yield strength. It is assumed that the yield strength is well modeled as a normally distributed random variable, but that the parameters of the normal distribution are unknown. Yield strength tests can be performed, thus sampling the distribution at a cost c per test.

In this experiment, each yield strength test represents one sample from the true material strength distribution, a normal distribution whose parameters are unknown to the designers. Specifically, the material strength is a random variable X such that:

$$X \sim N(\mu, \sigma^2) \tag{9.1}$$

The mean μ and variance σ^2 are unknown, and the goal of the information collection is to determine these parameters such that a good design decision can be made. The experiment consists of drawing a set of n samples $\{x_i\}_{i=1}^n$ from X. Each sample x_i that is drawn from the distribution is a piece of information that can be used to help characterize the true nature of the uncertainty. Unless the designers have infinite resources, they cannot collect the infinite number of samples necessary for a perfect characterization of the distribution. Instead, they need to determine when to stop collecting information—in this case, data samples.

As a designer collects data samples x_i, the marginal benefit of an additional sample decreases. For example, if the designer has only 10 samples, an 11th sample will usually be quite valuable; in contrast, if the designer has 1000 samples, the 1001st sample usually will be considerably less valuable, because the situations in which it significantly affects the estimates of the truth are small; it is only one of 1001 rather than one of 11 samples, and hence has much less impact. In this sense, information displays diminishing returns. At some point, the cost of gathering additional information will outweigh the benefit. Thus, the value of a sample is not merely inherent in the sample; rather, the value is measured as viewed from the perspective of the designer. A fundamental principle of information economics is that a decision maker (DM) should continue to collect information only as long as there is an information source available whose net value is positive. Putting the example problem into more standard micro-economic terms, a rational DM will stop taking data samples at the point where the marginal benefit of the next sample is less than or equal to the marginal cost of acquiring it. A formalization of the basic cost-benefit analysis noted above has been summarized in the context of information by Lawrence (1999). This work is summarized and expanded upon in the next section.

9.4 Mathematical problem formulation

In engineering design, the value of information can be measured by observing how the information affects the design decision. In this section and the next, the basic principles of information economics are explained in a general context.

9.4.1 Specifying probabilities over the state space

The set of all possible states of the world form a state space $X = \{x\}$. In the example problem, the state of the world is the actual material strength x of the material used in a particular pressure vessel. The material strength, or state, is assumed to be normally distributed with associated probability density function $p(x)$, with parameters that are unknown to the designer. The state of the world is outside the DM's control, so the DM can at best estimate the probabilities, thus forming the estimated distribution $\tilde{p}(x)$. As described in Chapter 3, these probabilities are interpreted according to a subjective interpretation of probability.

9.4.2 The payoff of a decision

As described in Section 3.6, for every decision problem a DM has a set of available actions $A = \{a_1, ..., a_m\}$ from which to choose one. The state of the world is described by a vector of uncertain quantities $s = \{{}_1x, ..., {}_nx\}$. Associated with every action-state pair was a particular set of consequences $\vec{g}(a,s) = \{g_1(a,s), ..., g_h(a,s)\}$.

In this chapter, the model is refined slightly in order to match the notation commonly used in the economic value of information literature (e.g. (Lawrence 1999)). It is assumed that there is only one uncertain quantity, and hence $s = \{{}_1x\}$ and the single variable x can be used to capture the state. It is also assumed that the only consequence of interest is the payoff $\pi(a,x)$ of an action a given the state of the world x, and thus $\vec{g}(a,s) = \{g_1(a,s)\} = \pi(a,x)$.

In the example problem, the action a consists of a set of design variables that specify the pressure vessel dimensions. The payoff function used in the example problem, shown in Equation (9.2), is highly skewed—the payoff when the vessel fails is largely negative (minus \$1 million), yet the payoff when it succeeds is only slightly positive (the selling price of \$200 minus the cost of the material used to build the pressure vessel). Skewed payoff functions are common in applications involving risk where rarely occurring events with severe consequences play a significant role in decisions. Note that for a given yield strength and design, the failure cost will either be zero (no failure) or a constant (the cost of the damage, lost productivity, etc. when the pressure vessel fails).

$$\begin{aligned}
&\pi(a,x) = P_{selling} - C_{material} * volume(a) - C_{failure} * \delta(a,x),\\
&\text{where:}\\
&\quad P_{selling} \equiv \text{selling price} = \$200\\
&\quad C_{material} \equiv \text{material cost per volume} = \$8500/\text{m}^3\\
&\quad x \equiv \text{true yield strength of pressure vessel}\\
&\quad a \equiv \text{design variables (radius, thickness, length)}\\
&\quad C_{failure} \equiv \text{cost incurred if vessel fails} = \$1,000,000\\
&\quad \delta(a,x) \equiv \text{failure indicator} = \begin{cases} 0 \text{ if } x \geq \sigma_{\max}(a) \\ 1 \text{ otherwise} \end{cases}
\end{aligned} \tag{9.2}$$

Direct use of the payoff function in the decision implies that the DM is risk neutral. If the DM is risk-averse or risk-taking, the payoff function should be mapped to a utility function according to this risk attitude. It also assumes that all of the DM's preferences are measurable in terms of the single attribute of payoff. The information economic approach presented in this paper can be used in such situations by performing the same cost-benefit analysis in terms of utilities instead of dollars.

The choice of a precise payoff function assumes perfect models of price, cost, and demand, models that do not typically exist. Imprecise value models could be used; however, this additional imprecision would translate into larger (less precise) bounds on the value of information. In this chapter, a precise value model is chosen in order to limit the number of sources of imprecision to one (the material strength characterization). Limiting the sources of imprecision allows for a clearer presentation of this new approach.

9.4.3 Making an optimal decision

Because of uncertainty in the state of the world x, the DM cannot know the payoff of any action with certainty. The DM seeks to maximize the expected payoff, given by $E_x[\pi(a,x)]$. The expectation is taken over all states x because that is what the DM is modeling as random. Note that ideally the expectation is taken with respect to the true distribution $p(x)$. However, in the example (and in most real world design scenarios), the DM does not know the true distribution $p(x)$, and must instead use his or her best-guess distribution $\tilde{p}(x)$. The DM thus makes an optimal decision, a^* such that

$$a^* = \arg\max_{\mathrm{a}} \ (E_{\tilde{p}(x)}[\pi(a,x)]). \tag{9.3}$$

The slight deviation from standard notation (i.e. writing $E_{\tilde{p}(x)}$ instead of E_x) is made to emphasize that the DM maximizes the expectation, as calculated using his or her subjective probability density function $\tilde{p}(x)$. A similar distinction must be made when determining the payoff of the decision. The true expected payoff is calculated using the true $p(x)$ that is unknown to the designer and is denoted as:

$$\pi_{true} = E_{p(x)}[\pi(a,x)]. \tag{9.4}$$

The estimated expected payoff according to the designer's best-guess distribution is:

$$\pi_{\tilde{p}(x)} = E_{\tilde{p}(x)}[\pi(a,x)]. \tag{9.5}$$

This payoff $\pi_{\tilde{p}(x)}$ will in general differ from the true payoff π_{true}. Although Lawrence (Lawrence 1999) does not make this distinction in his work, the distinction is crucial in cases in which the designer has only imprecise information about the random distribution.

9.4.4 Information and information sources

The definition of information varies significantly by subject and application. In this dissertation, Lawrence's definition (Lawrence 1999) is modified slightly and information is defined as any stimulus that changes the recipient's best-guess probability distribution $\tilde{p}(x)$ over a well-described set of states, $X = \{x\}$.

An information source is anything that provides information. This information arrives in the form of a message y taken from a set Y_I that information source I can deliver. There is some probability $p(y)$ associated with receiving a particular message y from the information source.

In the pressure vessel design problem, the information source is the yield strength testing process, and a message is the result of a single yield strength test—that is, one observation of material strength. Information economics studies whether it is valuable to pay an information source for a message. Before the message is received, a DM does not know what information that message contains, and therefore the DM does not know exactly how it will change his or her subjective probability distribution $\tilde{p}(x)$ over the state space. In turn, the DM does not know how the message will affect the decision a^* and its payoff. Thus, a DM should apply the principles of information economics to arrive at a formal metric for determining if the benefit of a message outweighs the cost of acquiring it—the value of information.

9.4.5 The value of information

Two possible decision are now considered: the first decision is made using the current state of information, and the other is made after the receipt of message y. In the first case, assume the DM's subjective probability distribution of the states is represented as $\tilde{p}(x)$. These are the prior probabilities, and the optimal prior decision a_0^* is given by

$$a_0^* = \arg\max_a \; (E_{\tilde{p}(x)}[\pi(a,x)]) \tag{9.6}$$

After the message y is received and incorporated into the DM's knowledge, the DM has an updated posterior probability distribution $\tilde{p}(x \mid y)$. The corresponding optimal decision a_y^* is given by

$$a_y^* = \arg\max_a \left(E_{\tilde{p}(x|y)}[\pi(a,x)]\right) \tag{9.7}$$

How can these two decisions be compared? If the comparison is made after the true state of the world x is revealed, the ex-post gross value of the message y—where gross implies before factoring in cost—can be calculated for the particular realized state x as:

$$\upsilon\left(y \mid x\right) = \pi(a_y^*, x) - \pi(a_0^*, x) \tag{9.8}$$

This represents the amount that the receipt of message y (and the incorporation of its information into the decision) changed the decision maker's payoff, given the particular outcome x of the state.

The term value is used throughout this paper in a marginal sense, that is, in terms of differences. The ex-post gross value of a message y is the marginal payoff of acquiring that message—the difference between the payoff of the decision with and without the information from message y. This gross value can be positive, negative, or zero. It will be positive if the message leads the DM to chose an action a_y^* that has a higher payoff under realized state x than action a_0^*. It will be negative if the message in someway misled the DM into choosing an action a_y^* that has a lower payoff than the prior decision a_0^*. If the message did not change the choice of action, such that a_y^* is the same as a_0^*, then the ex-post gross value is zero.

The previously defined ex-post gross value is not useful for determining the potential change in payoff of receiving a message because it measures the actual benefit, which can only be known after the decision is made and the truth realized. It is common knowledge that a good decision can lead to a bad outcome, especially if a rare, adverse state of the world is realized—a situation referred to in the vernacular as bad luck. Conversely, a bad decision can lead to a good outcome—a case of good luck.

Rather than assessing the value of a message for a particular state x, a DM is really interested in the expected value over all the possible states of the world. The gross value of a message y is defined as the expected difference in the payoff with and without the message, such that:

$$\upsilon(y) = \text{gross value}(y) = E_x[\pi(a_y^*, x) - \pi(a_0^*, x)] \tag{9.9}$$

Calculating the true gross value of a message requires the expectation over the true distribution $p(x)$, which is not available to the DM.

To complicate matters further, Equation (9.9) is valid for analysis of the value of a particular message y only after it is received. However, when the DM needs to decide whether or not to purchase a message, the content of the message—that is the particular message y from the set Y_I of all possible messages from the source I—is also unknown. When purchasing a message, it is as if the DM is purchasing a sealed envelope; he or she does not know what is inside until after buying and opening the envelope. The DM must therefore consider the value of the information source I instead of the value of a single message.

If the DM had access to the true probability distribution of the messages, $p(y)$, over the set Y_I, he or she could calculate the gross value of the next message from an information source I:

$$\text{gross value}(I) = E_y E_x[\pi(a_y^*, x) - \pi(a_0^*, x)] \tag{9.10}$$

Because the DM does not have access to the parameters describing the true probability distribution of the messages $p(y)$ or of the states $p(x)$, Equation (9.10) cannot be used directly to estimate the value of an information source. In this chapter, an approach for bounding the value of information that incorporates the imprecision of the DM's information state is developed.

A final definition that ties the notion of value back to the fundamental concept of cost-benefit analysis in information economics is net value. A message y must be purchased at some cost; resources need to be expended in order to acquire more information. Denoting this as $\text{cost}(y)$, the net value of a message is defined as

$$\text{net value}\left(y\right) = E_x[\pi(a_y^*, x) - \pi(a_0^*, x)] - \text{cost}\left(y\right). \tag{9.11}$$

Similarly the net value of the next message from an information source is:

$$\text{net value}(I) = E_y E_x[\pi(a_y^*, x) - \pi(a_0^*, x)] - \text{cost}(I), \tag{9.12}$$

where $\text{cost}(I)$ is the cost of receiving one message y from information source I.

If the DM's goal of making a cost-benefit tradeoff during information collection, is revisited, the information economic principle can now be stated that a designer should purchase a message from an information source if and only if the net value of that information is positive. According to Equation (9.12), this requires the calculation of expectations across the distributions $p(x)$ and $p(y)$, which in general are not known to a designer. Attention will be returned in Section 9.6 to the problem of not knowing $p(x)$ and $p(y)$ after illustrating the simpler case of known probabilities.

9.5 Example with known probabilities

In this section, an example is presented to illustrate the calculation of value of information in the hypothetical case of known probabilities. This example is extended to the more practical case of unknown probability parameters in Section 9.6. While the information used in this example is not available to a DM, it is useful for illustrating the basic approach, shown in Figure 9.1.

It is assumed that there is an omniscient supervisor overseeing the experiment. This supervisor knows the true distribution and can perform the actions shown in the gray boxes. These actions are normally not available to the DM and are used in the experiment only for meta-analysis, not in the actual design decision. In this approach, the DM begins with the observed set of samples $\Sigma = \{x_i\}_{i=1}^{n}$. The goal is to determine whether it is valuable to collect an $(n+1)^{st}$ sample given the existing n samples. The DM first uses this set of samples to construct a best-fit distribution $\tilde{p}(x)$, and then to choose an optimal design a_0^*, as shown on the left side of the figure. The DM then receives a hypothetical additional sample y_j from the supervisor. The DM constructs a new best-fit distribution $\tilde{p}(x \mid y_j)$ and makes a new decision $a_{y_j}^*$. The difference in expected payoffs of the two decisions is then calculated by the supervisor to determine the true expected gross value $\upsilon(y_j)$ of the particular message y_j. This process of calculating the value of an additional sample is repeated over many y_j to calculate the average value of the next sample for a particular starting set of n samples, which will be denoted as $V(n+1 \mid \Sigma)$.

Recall that the net value of the next piece of information depends on the prior decision a_0^*, which in turn depends on the existing data samples. For example, the net value of purchasing an 11th sample from the information source will depend on the first 10 samples. If the initial 10 samples just happen—by chance—to yield very good estimates of the distribution parameters, then the net value of the 11th sample will be small, but if they yield bad estimates of the distribution parameters, then the net value of the 11th sample could be large. Consequently, the next step is to repeat the experiment over many initial sample sets Σ, which gives the average gross value of the next sample, denoted $V(n+1)$.

The final step of the experiment is to repeat the process for different initial sample sizes. By repeating the calculation over many initial sample sizes, a curve can be constructed for the average net value of an additional sample at different sample sizes, as

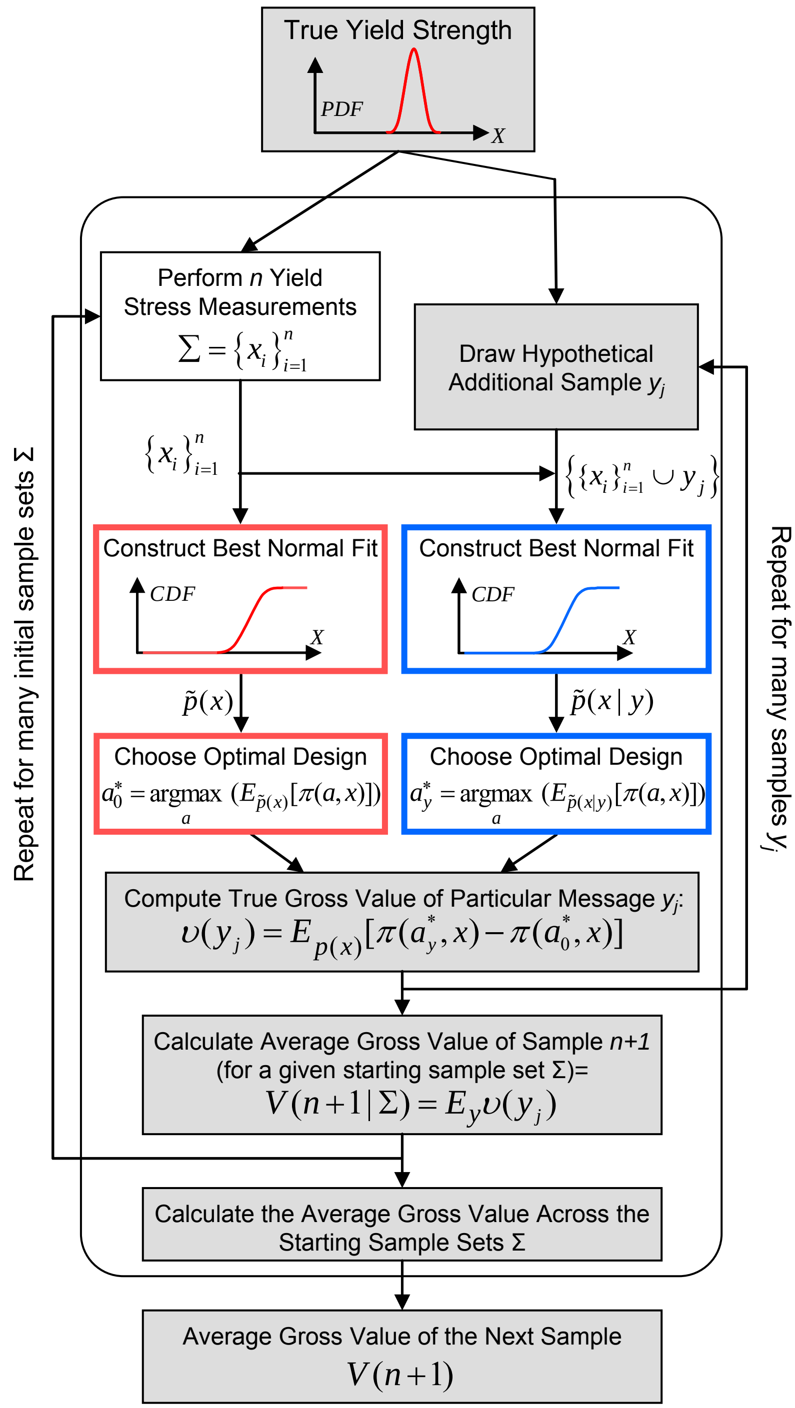


Figure 9.1. Calculating the value of information with known probabilities

shown in Figure 9.2. This figure can be interpreted, for example, as follows: at a prior sample size n=32, the average net value of an additional sample (the 33rd sample in this case), is about \$2. The net value of the 52nd sample, starting from 51 samples, is negative, but the net value of the 51st sample is positive. This means that the 52nd sample is the first sample whose average net value is negative; therefore, stopping at 51 samples will result, on average, in the highest expected utility. Note that this conclusion is drawn using the true $p(y)$ and $p(x)$, which are not available to the DM.

The results can also be interpreted by considering the net expected payoff, which is the payoff of the design that would have been realized if no additional information were collected, less the cost of the already collected n samples:

$$\text{net expected payoff} = E_{p(x)}[\pi(a^*, x)] - n \cdot \text{cost}(y) \tag{9.13}$$

The results are shown for different sample sizes in Figure 9.3. Again, because the actual observed samples affect the payoff, the payoff of the design is averaged over many initial sample sets. The relationship between this result and the net value of additional samples should be clear; the maximum net expected payoff occurs at the same sample size at which the net value of an additional sample first becomes negative. Recalling that

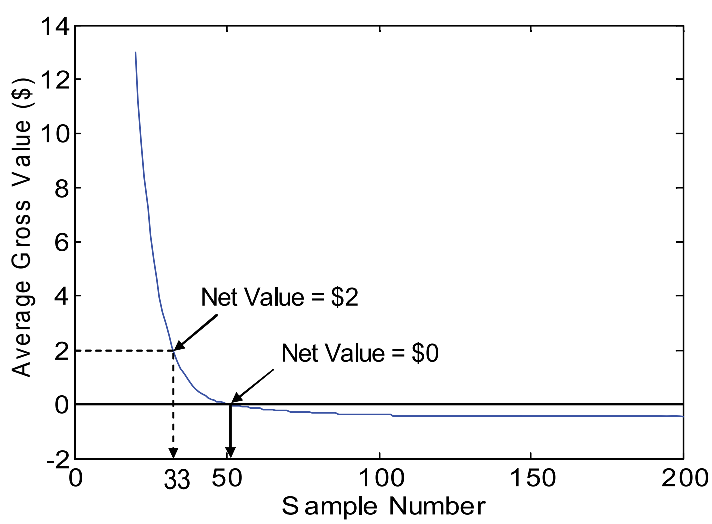


Figure 9.2. Net gain in payoff per sample

the net value is defined in a marginal sense, moving from 51 samples to 52 samples means a decrease in total payoff of the decision, as is revealed in both plots.

In the preceding analysis, it appears simple to determine the optimal number of samples to collect. However, this simplicity is due to the omniscient supervisor having precise knowledge of the true distributions $p(y)$ and $p(x)$. In the example problem the information source is an unbiased model of the truth, which means that $p(y)$ and $p(x)$ are identical yet unknown—they both describe the unknown true material strength. The characterization of $p(x)$ is the DM's indirect goal for data collection—the DM wants to characterize $p(x)$ well enough that the design based on the estimated $\tilde{p}(x)$ is acceptable.

To determine the value of information during the actual design process, the DM needs a method by which he or she can estimate the net value of an additional data sample when the parameters describing $p(y)$ and $p(x)$ are unknown. A new approach is proposed that uses imprecise probabilities to calculate an interval of net value for an additional sample.

What performance characteristics should be expected or demanded of this approach? Insight can be gained by examining the distribution of the net payoffs about

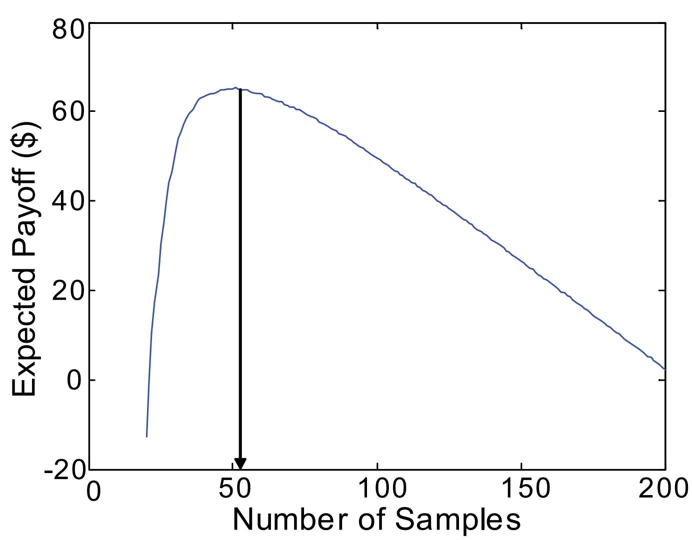


Figure 9.3. Net expected payoff of the design

the expected value curve of Figure 9.3. Box plots for sample sizes of 50, 100, and 150 are shown in Figure 9.4. The plots are constructed with the whiskers at the 0.1% and 99.9% quantiles, and the boxes from 25% to 75%. The extreme skewness of the box-plots is due to the skewed payoff function; that is, the cost of slightly under-designing the pressure vessel is large compared to the cost of slightly over-designing it.

The box plots reveal that both the variance of the payoff and the chance of a catastrophic result decrease as the sample size increases but that, simultaneously, the expected net value decreases significantly. The behavior shown in Figure 9.4 suggests that a reasonable estimation of the optimal number of samples (when the DM has only imprecise knowledge about the true distribution) will often be well beyond 51 (the optimal stopping point based on expected value), because by stopping at 51 samples, a DM still faces a very large downside risk. It is important to consider the distributions in Figure 9.3 and Figure 9.4 when developing an approach for determining the value of additional samples; however, in practice, an engineer will not have this information available for decision making, as considered in the next section.

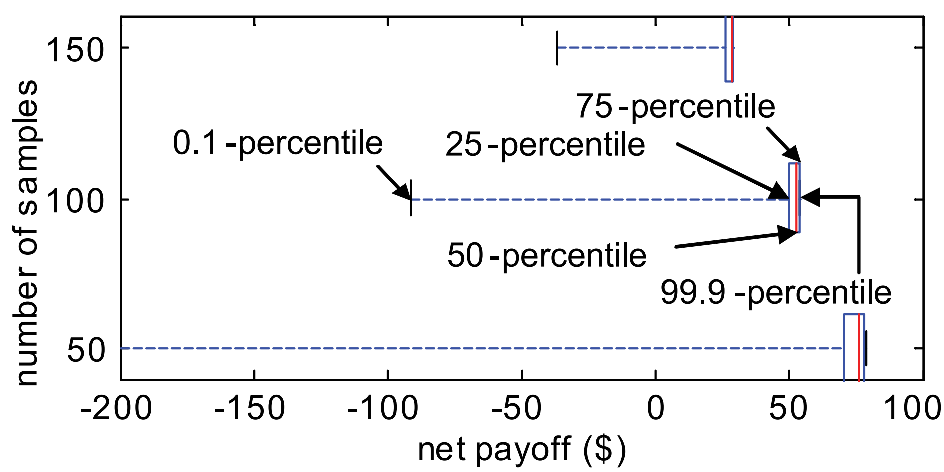


Figure 9.4. Box plots for various sample sizes

9.6 Estimating the value of information

In this section, a new approach to bounding the gross value of the next message from an information source is developed. Section 9.6.1 is a description of the decision policies used in the example. In Section 9.6.2, the use of imprecise probabilities is motivated. The specific approach for estimating the value of information is explained in Section 9.6.3. Finally, the computational experiment is described in Section 9.6.4.

9.6.1 Design decision policy

According to Equation (9.3), the DM chooses the design action that maximizes the expected payoff, with the expectation calculated using $\tilde{p}(x)$. This distribution is derived by assuming that the material strength is normally distributed and then using the sample mean and sample variance of the observed samples as precise estimates of the true mean and variance. Other work (see Chapter 9) has presented a decision policy that incorporates imprecision into $\tilde{p}(x)$ during the solution phase of the design decision, much as the approach in this paper incorporates imprecision into the problem formulation phase. Nevertheless, in this chapter a decision policy based on a best-fit distribution is used in the problem solution phase in order to isolate the effect of accounting for imprecision in the problem formulation phase—that is, to emphasize the contributions of applying information economics to decisions regarding the collection of information to support the actual design decisions. A noted item for future work is the combination of these approaches into one unified approach that explicitly considers imprecision throughout the design process.

9.6.2 Motivation for using imprecise probabilities

One motivation for using imprecise probabilities to represent the DM's state of information is that the use of precise probabilities does not enable useful estimates of value. The necessity of an alternative to precise probabilities is illustrated in the following example. Assume that the DM represents his or her state of information $\tilde{p}(x)$

precisely. Using this information, the DM chooses an optimal design a_0^* according to Equation (9.6), using $\tilde{p}(x)$ when evaluating the expectation.

Now assume that the DM acquired an additional data sample y. With this information, the DM can create a new subjective distribution $\tilde{p}(x \mid y)$, where in general $\tilde{p}(x \mid y) \neq \tilde{p}(x)$. The DM would then choose an optimal design a_y^* according to Equation (9.7), using $\tilde{p}(x \mid y)$ when evaluating the expectation.

If the DM wanted to calculate the gross value of this message y, he or she would use Equation (9.9), repeated here for clarity:

$$\upsilon(y) = \text{gross value}(y) = E_x[\pi(a_y^*, x) - \pi(a_0^*, x)]\,. \tag{9.9}$$

Ideally the expectation E_x would be taken over the true $p(x)$, but the parameters of this distribution are unknown. The DM's two best options for approximating $p(x)$ are $\tilde{p}(x)$ and $\tilde{p}(x \mid y)$.

If the DM uses $\tilde{p}(x)$ as the best estimate of $p(x)$, Equation (9.9) can be rewritten:

$$\upsilon(y) = E_{\tilde{p}(x)}[\pi(a_y^*, x) - \pi(a_0^*, x)]\,. \tag{9.14}$$

or, by distributing the expectation as:

$$\upsilon(y) = E_{\tilde{p}(x)}[\pi(a_y^*, x)] - E_{\tilde{p}(x)}[\pi(a_0^*, x)]\,. \tag{9.15}$$

According to Equation (9.6), the design decision a_0^* maximizes $E_{\tilde{p}(x)}[\pi(a, x)]$, thus

$$E_{\tilde{p}(x)}[\pi(a_y^*, x)] \leq E_{\tilde{p}(x)}[\pi(a_0^*, x)] \tag{9.16}$$

This means that the gross value of message y is always estimated to be zero or negative, no matter how much new information is available. Yet intuitively, the gross value of additional information should often be positive—acquiring information should improve the DM's ability to make a good decision on average. Consequently, it is not useful to make information collection decisions using $\tilde{p}(x)$.

If the DM instead used the posterior distribution $\tilde{p}(x \mid y)$, Equation (9.9) can be rewritten as:

$$\upsilon(y) = E_{\tilde{p}(x|y)}[\pi(a_y^*, x) - \pi(a_0^*, x)]. \tag{9.17}$$

Expanding the expectation yields:

$$\upsilon(y) = E_{\tilde{p}(x|y)}[\pi(a_y^*, x)] - E_{\tilde{p}(x|y)}[\pi(a_0^*, x)]. \tag{9.18}$$

According to Equation (9.7), design decision a_y^* maximizes $E_{\tilde{p}(x|y)}[\pi(a, x)]$, thus:

$$E_{\tilde{p}(x|y)}[\pi(a_y^*, x)] \geq E_{\tilde{p}(x|y)}[\pi(a_0^*, x)]. \tag{9.19}$$

In this case, the gross value is always calculated to be positive or zero, which is also unreasonable; there will always be "unlucky" samples, or messages, that lead to a worse design. Another objection to using the precise $\tilde{p}(x \mid y)$ is that it has no use in decision making, because $\tilde{p}(x \mid y)$ is only available after the information message y is collected.

This exercise illustrates that an information collection policy based upon the assumption of precisely characterized knowledge about the true distributions is impractical. However, the principles of information economics can be implemented using an approach based on imprecise probabilities that provides useful bounds on the value of information, as described in the next section.

9.6.3 Bounding the value of information

An overview of the approach for bounding the value of future information collection is shown in Figure 9.5. The DM begins with the actually observed set of data samples $\Sigma = \{x_i\}_{i=1}^n$. The DM first uses this sample to construct a best-fit normal distribution and to choose an optimal design a_0^*—the left side of the figure. The DM then uses the observed samples to construct a parameterized p-box using the method described in Section 4.3. This method creates the p-box by using confidence intervals on the mean and variance given by Equations (9.20) and (9.21) (Hines, et al. 2003), where α is the confidence level, n is the number of samples and s is the sample standard deviation, $t_{\alpha/2,n-1}$ is the t-statistic and $\chi^2_{\alpha/2,n-1}$ the Chi-squared statistic.

$$[\underline{\mu}, \overline{\mu}] = [\hat{\mu} - t_{\alpha/2,n-1} \frac{s}{\sqrt{n}}, \hat{\mu} + t_{\alpha/2,n-1} \frac{s}{\sqrt{n}}] \tag{9.20}$$

$$\left[\underline{\sigma}^2, \overline{\sigma}^2\right] = \left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right] \tag{9.21}$$

Recall that by assumption, this model of information—the p-box—contains the CDF that corresponds to the true distribution $p(x)$. The DM discretizes the p-box, as described below, and selects a single normal distribution from the p-box to represent both $\tilde{p}_i(x)$ and $\tilde{p}_i(y)$. Because the information source is an unbiased model of the truth in the example problem, these two distributions are identical; they both describe the unknown true material strength, meaning both x and y are realizations of the same random process. This selected distribution is used to estimate the gross value of collecting an additional piece of information through the use of Equation (9.10) with $p(x) = \tilde{p}_i(x)$ and $p(y) = \tilde{p}_i(y)$. If the DM repeated this for every normal distribution inside the p-box, one of the calculated values would be the true gross value of the next piece of information. Clearly, the DM cannot try every distribution, so the following procedure is proposed.

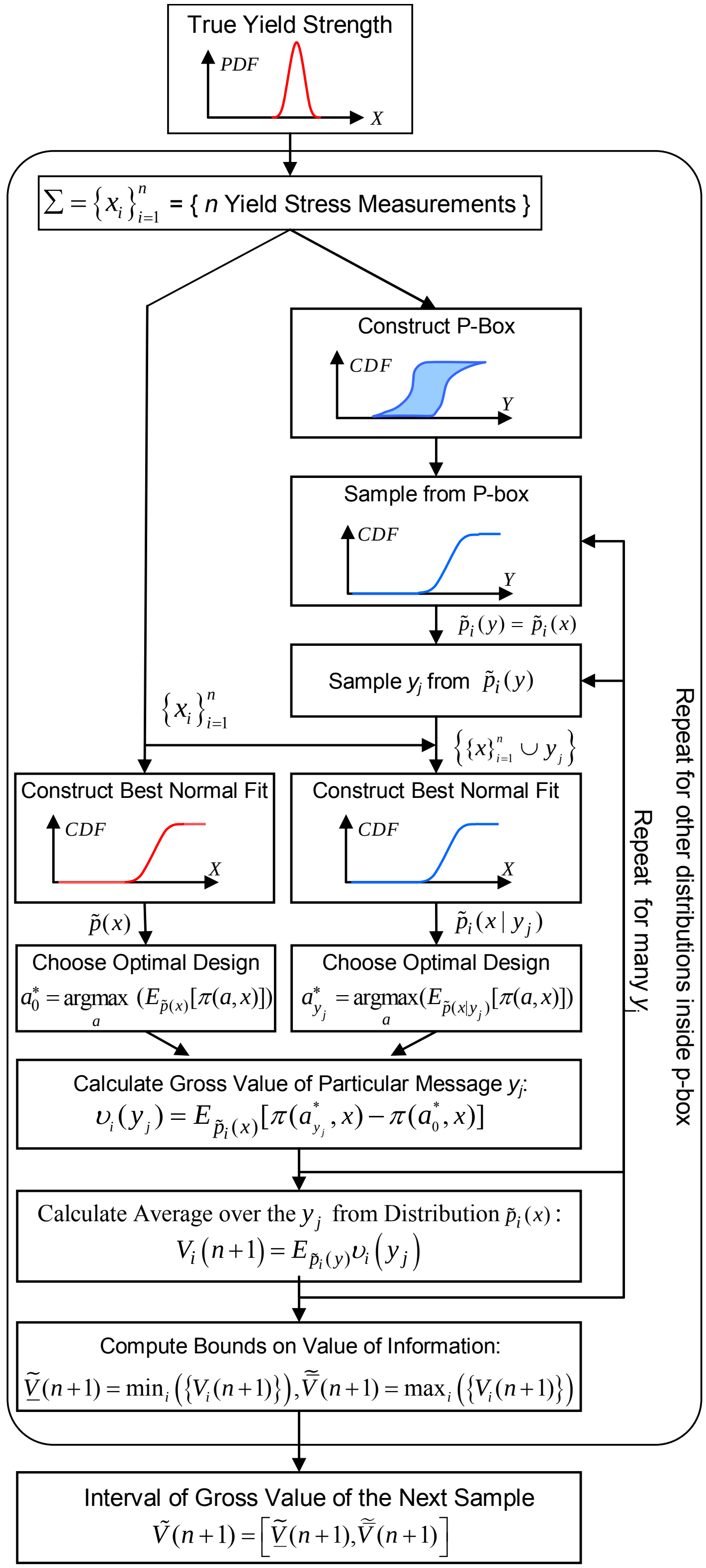


Figure 9.5. Overview of approach using imprecise probabilities to bound the value of information

The DM can partition the p-box into a finite, representative set of distributions. This is done by discretizing the confidence intervals on the mean and variance. The DM pairs all the combinations of mean and variance, resulting in a set of distributions such as shown in Figure 9.6. Future work will explore more efficient methods for this partitioning such as concepts from design of experiments, direct manipulation of the p-boxes, random sampling across the confidence intervals, or the optimized methods developed by Bruns and co-authors (Bruns 2006, Bruns and Paredis 2006, Bruns, et al. 2006). For illustration of the concept of the method for bounding the value of future information collection, the simple computational method suffices.

The DM selects one distribution, say $\tilde{p}_i\left(x\right)$, from this finite set and assumes that this distribution is the true distribution $\left(\tilde{p}_i\left(x\right)=p(x)\right)$. The DM then calculates the gross value of taking the $(n+1)^{st}$ sample, denoted $V_i(n+1)$, via a Monte Carlo simulation, as follows.

Given the assumed message distribution, $\tilde{p}_i\left(y\right)=\tilde{p}_i\left(x\right)$, the DM can draw a

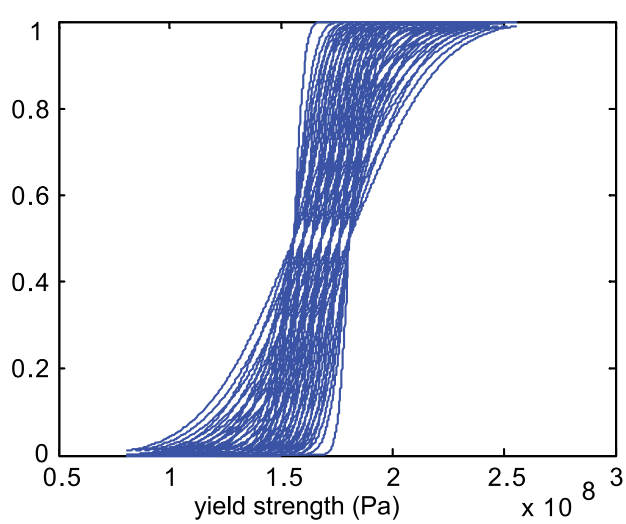


Figure 9.6. Various distributions in the P-box

hypothetical next sample, y_j, from the distribution. This sample is used, along with the actually observed samples $\{x_i\}_{i=1}^{n}$, to estimate a new posterior distribution $\tilde{p}_i(x \mid y_j)$. The DM uses this distribution to choose the optimal design, $a_{y_j}^{*}$, for the given distribution and hypothetical sample. The DM then evaluates the expected payoff of this design using the assumed $\tilde{p}_i(x)$, and calculates the gross value $\upsilon_i(y_j)$ of that particular y_j. The DM repeats this for many different y_j's drawn from $\tilde{p}_i(y)$ and calculates the average, or expected, gross value of the next message with distribution $\tilde{p}_i(x)$ assumed to be the true distribution, denoted as $V_i(n+1)$. Finally, the DM repeats this process for all $\tilde{p}_i(x)$ in the chosen set (from the parameterized p-box). This results in a set of gross values $\{V_i(n+1)\}$.

Recall that if the p-box contains the true distribution[17] and if it were sampled exhaustively, then one of the values $V_i(n+1)$ in this set would be the true gross value of the $(n+1)^{st}$ sample, given the previously observed n samples. The set $\{V_i(n+1)\}$ would then form an interval $V(n+1)=[\underline{V}(n+1),\overline{V}(n+1)]$. In this approach, the p-box is only finitely sampled and it may not contain the true distribution, so the set of values $\{V_i(n+1)\}$ only gives an approximate interval, $\tilde{V}(n+1)=[\underset{\sim}{\underline{V}}(n+1),\overset{\approx}{V}(n+1)]$, with the lower and upper-bounds defined as $\underset{\sim}{\underline{V}}(n+1)=\min_i(\{V_i(n+1)\})$ and $\overset{\approx}{V}(n+1)=\max_i(\{V_i(n+1)\})$. The accuracy of these estimated intervals improves as the density of sampling from the p-box increases.

Based on this interval of value for the next sample, the DM decides if another sample should be taken. If another sample is taken, the process repeats itself starting with the larger set of $n+1$ data samples $\Sigma=\{x_i\}_{i=1}^{n+1}$. It should be noted that in general the

[17] The p-box box may not contain the true distribution because it is created using confidence intervals at less than the 100% confidence level because the 100% level in general yields infinite bounds.

p-box and hence the discretized distributions used in the analysis will be different for this new data set, meaning that in general the quantities $\tilde{p}_i\left(y \mid n\right) = \tilde{p}_i\left(x \mid n\right)$ are not equal to the quantities $\tilde{p}_i\left(y \mid n+1\right) = \tilde{p}_i\left(x \mid n+1\right)$.

9.6.4 Computational Experiment

The method defined in the previous sub-section will now be applied to the pressure vessel design problem described in Sections 5.2 and 9.3. The experiment proceeds according to the approach shown in Figure 9.5 and is repeated for sample sizes up to 200. This generates intervals on the gross value for one particular sequence of random samples $\{x_i\}$. This experiment is then repeated many times to generate multiple sample traces, meaning multiples sets of random samples.

9.7 Results

Using the approach and experiment described above, bounds on the gross value of the next piece of information can be found. A graph of these bounds ($\tilde{V}(n+1) = [\underline{\tilde{V}}(n+1), \overline{\tilde{V}}(n+1)]$) for a particular sequence of samples $\{x_i\}_{i=1}^{n}$—a particular sample trace—is shown in Figure 9.7. A trace represents the bounds on the gross value of the next sample for a set that grows every time a sample is acquired. For example, if the set of samples $\{x_i\}_{i=1}^{10}$ has already been collected, the bounds of the value of the 11$^{\text{th}}$ sample can be estimated. If the 11$^{\text{th}}$ sample (x_{11}) is actually acquired, then a new set is formed of the existing samples plus the new sample, such that $\{x_i\}_{i=1}^{11} = x_{11} \cup \{x_i\}_{i=1}^{10}$. Figure 9.8 shows the upper-bound, lower-bound, and midpoint for two additional traces in the vicinity of their crossing of the cost line—the zero net value point. The curves in the two figures reveal several interesting characteristics, as discussed in the following sections.

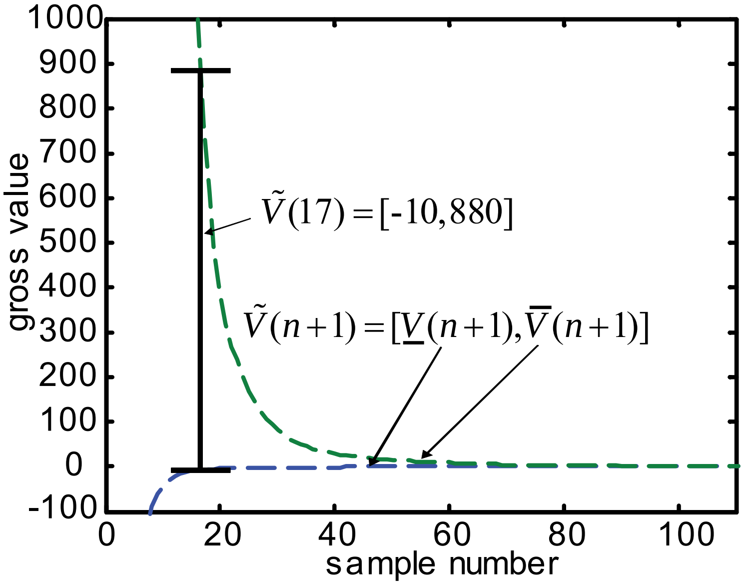


Figure 9.7. Example high-level behavior of gross value

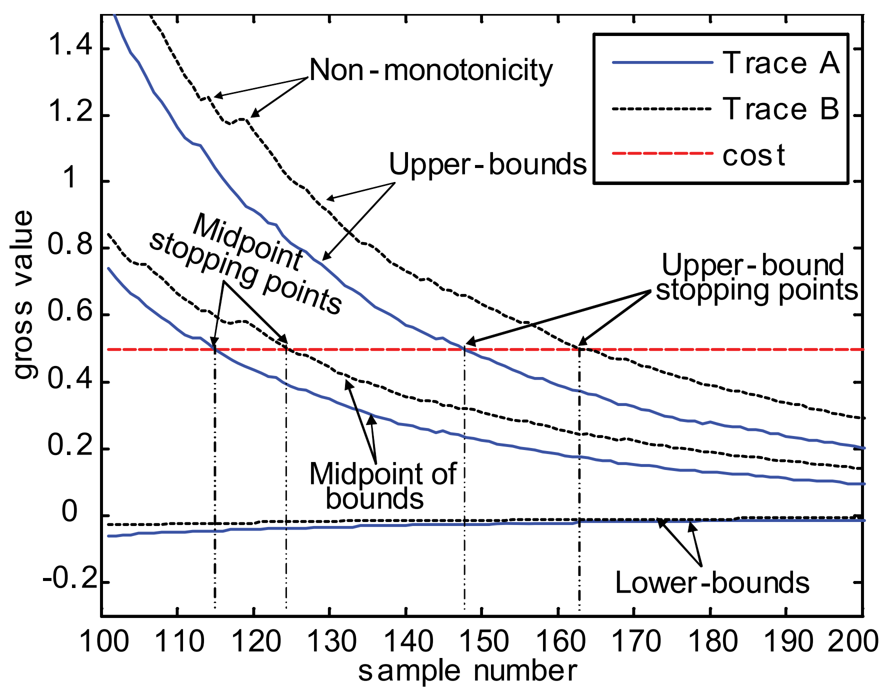


Figure 9.8. Two example traces of gross value

9.7.1 Small sample sizes yield large value intervals

Examination of Figure 9.7 reveals that the potential value of the next sample for small sample sizes covers a very large range that is skewed towards the positive side. For example, the gross value of the 17th sample is in the interval [-10, 880]. Based on a strict interpretation of traditional decision policies, an interval bounding zero leads to indeterminacy. However, in the context of information collection, a decision musts be made. This can be understood as follows. In an information collect decision, there are two alternatives—collect information or do not collect information. These actions are exhaustive of the universe; there are no other actions possible. Normally when information is indeterminate, no decision is made. In an information collection decision, not making a decision is equivalent to deciding not to collect information.

Given this situation, essentially a decision must be made, and a decision policy that can resolve the indeterminacy is required. In this dissertation, one extreme is assumed—the DM stops collecting data when the upper-bound on the gross value is less than the cost—that is, when the upper-bound on the net value is negative. This is a so-called Γ-maximax policy (Berger 1985, Schervish, et al. 2003). At the accepted confidence level, the true value is assumed to lie in the interval, so this represents the point at which the true net value cannot exceed zero, therefore, no rational DM would take an additional sample. Other possible policies for managing interval-based decisions include maximality (Walley 1991), Γ-maximin (Berger 1985), E-admissibility (Schervish, et al. 2003), and the Hurwicz criterion (Arrow and Hurwicz 1972). The general use of these policies in engineering design is an area for future work.

9.7.2 The bounds on value are not monotonic

In a general sense, it is reasonable to expect that the value of additional samples will decrease as n increases. However, each trace represents one sequence of actually observed samples. Thus, the gross value of the i^{th} sample is estimated based on the first

$i-1$ samples. Once the i^{th} sample is collected, the value of the $(i+1)^{st}$ sample is calculated using all i acquired samples. If the actually acquired i^{th} sample is really "lucky" or "unlucky", the gross value of the next sample can change significantly, potentially yielding non-monotonic bounds. An example of such non-monotonicity is labeled in Figure 9.8. Non-monotonicity can result in multiple cost line crossings, but these crossings were never observed to be more than a few (3-5) sample sizes apart. Because the bounds are already estimates, a deviation of a few samples is not likely to be significant.

9.7.3 The lower-bound is always non-positive

It is worth noting that the lower-bound on the interval will always be non-positive, i.e., given the available information, it is always possible that the gross value of the next piece of information will be less than zero. This happens because the best-fit distribution $\tilde{p}(x)$ on which the design decision is based is always contained in the set of distribution samples from the p-box—it is a candidate for the truth in our approach. This means that during the calculation of the interval on gross value, this distribution will be considered as the truth at some point, yielding the situation described in Equation (9.14)—if the DM's estimate already is the true distribution, which is possible though rare, then no information can make the estimate any better; it will in fact often make the estimate worse because the next random sample will rarely be a perfect confirmation of the existing information.

9.7.4 Examining the net value

Another point of note is the relationship between the gross value and cost. In practice, there is a relationship between the number of pressure vessels being designed and the cost, because the cost of information collection is amortized over all the pressure vessels. In this example, it is assumed that each yield strength test on a material sample costs \$0.50 per pressure vessel, and the discussion is focused on the design of one pressure vessel. Other cost functions could be used without adding significant complexity. With the cost fixed at \$0.50, an experiment following Trace A and using the upper-bound decision rule will stop with the 147th sample, because the upper-bound on the gross value of the 148th sample is less than the cost, as can be seen in Figure 9.8. The same logic can be applied to Trace B. In this case, a DM would collect 162 samples because the upper-bound on the gross value of the 163rd sample is less than the cost; the net value is negative.

In this section, two representative results have been presented. The overall results consist of many sample traces that can be analyzed in the same way as the examples above. In the next section, the performance of the method is examined.

9.8 Comparison of realized payoffs

Using the true material strength distribution $p(x)$—which is not known by the designers—an omniscient supervisor could evaluate Equation (9.13) to determine the actual expected payoff of the optimal design, a^*, after each sample. The results of this evaluation for Trace A from Figure 9.8 are shown in Figure 9.9. Each point represents the true expected net payoff (y-axis) of a design chosen based on the current number of samples (x-axis). Figure 9.9 is similar in nature to Figure 9.3 in Section 9.5; the volatility of the curves in Figure 9.9 is due to the fact that the value along a single trace is investigated, rather than an average of the value of the next sample over many traces as was shown to Figure 9.3. Because a DM would never create all of these designs and does not have access to $p(x)$, this is a hypothetical exercise that only the omniscient supervisor can perform using the true distribution $p(x)$.

The results in Figure 9.9 indicate that, given the actual observed sequence of samples, the DM would have been best-off stopping earlier (at 5 samples) than the new approach suggests (at 147 samples). In this example, the DM loses about 60% of the payoff by collecting the additional 142 samples.

Is this a result of the stopping policy? Because the DM stops collecting information only when he or she is absolutely sure that the value of the next piece of information is less than its cost, the Γ-maximax decision policy will often be overly conservative. An alternative policy would be to use the midpoint of the bounds, a special case of the Hurwicz criterion (Arrow and Hurwicz 1972). Using this stopping rule the DM would collect 114 samples, for Trace A in Figure 9.9. This still results in a loss of 50% payoff from the optimal.

Is such a loss in payoff justified? In the discussion surrounding the distribution of payoffs and Figure 9.4, it is concluded that the DM may wish to go beyond the average "optimal" stopping point due to the imprecision in the DM's knowledge and the large downside risk of stopping too soon. The actual expected net payoffs for trace B from Figure 9.8 are shown in Figure 9.10. In this example trace, it turns out that given the actually observed samples, it would have been much worse to stop after 80 samples as

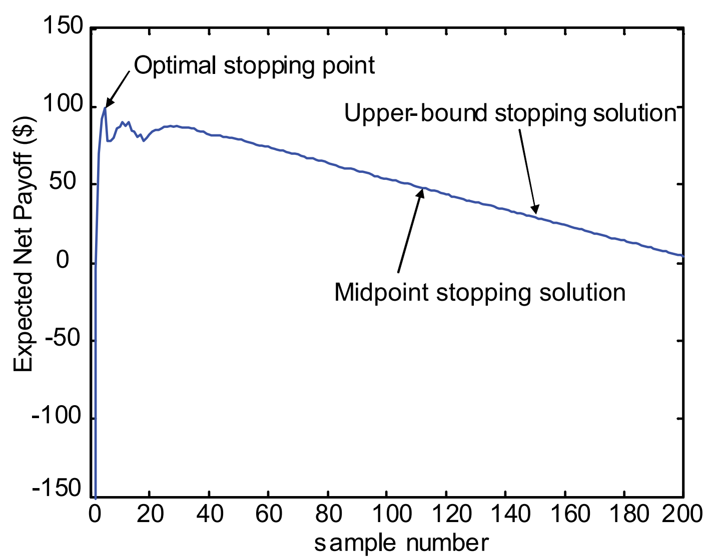


Figure 9.9. Actual expected net payoffs for Trace A

compared to 100. According to Figure 9.8, the midpoint stopping rule would have stopped at 124 samples for this trace. While this is still about 50% below the optimal, it yields a significantly better result than a policy that would have stopped at 80 samples.

Before ending this analysis, one last trace is presented in Figure 9.11. For this trace, the optimal stopping point would have been at 110 samples. The midpoint stopping rule would have led to the collection of 130 samples. This is relatively close to the optimal but still results in some loss of payoff. What causes the optimal stopping point to be so high in this case? One contributing factor is that the first five actual samples were 192 MPa, 200 MPa, 194 MPa, 197 MPa, and 181 MPa. These are all above the true mean of 180 MPa. This initial "unlucky" bias leads to a severe over-estimate of the material strength, which in turn leads to a severe under-design of the pressure vessel. Consequently, the pressure vessel fails much more often than expected, leading to a significantly higher average failure cost. This example indicates how sensitive the design can be to the sample data, and why a large number of samples may be needed to reach a stable result.

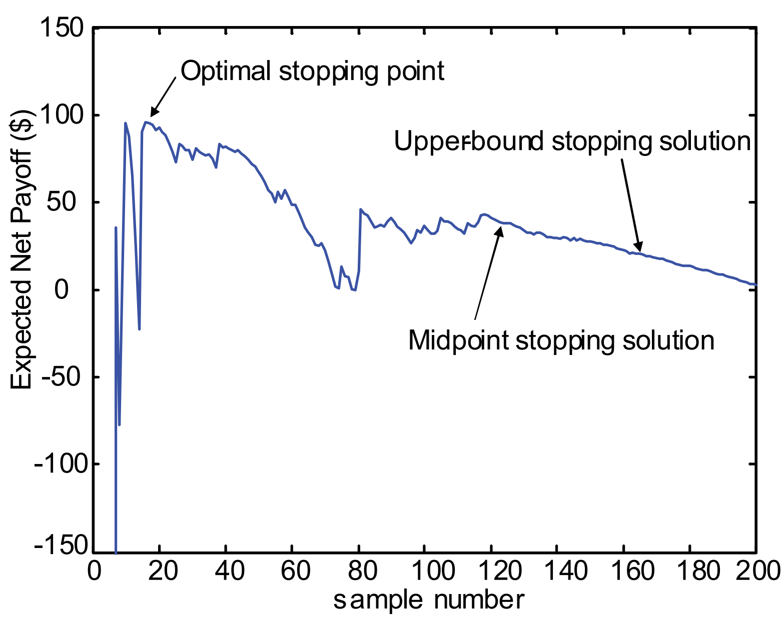


Figure 9.10. Actual expected net payoffs for Trace B

Before reaching a conclusion on the effectiveness of this approach of bounding the value of information, it is emphasized that the DM would not have the actual expected payoff curves available. Therefore, the DM does not know if he or she is in an example similar to that of Figure 9.9, Figure 9.10, Figure 9.11, or something else altogether. A conservative policy therefore leads the designer to keep taking samples until he or she is reasonably sure that there is no chance of a large negative payoff; that is, samples are taken until the downside risk is acceptable.

9.9 Future work

In this chapter, a foundation for applying information economics to engineering design decisions involving the collection of information about probability distributions with unknown parameters was laid, but there is still significant room for improvement and additional exploration.

Decision policies for gathering information. In this chapter, an approach for bounding the value of collecting additional data samples was developed. These bounds need to be resolved according to some policy in order to make a decision. Given just the

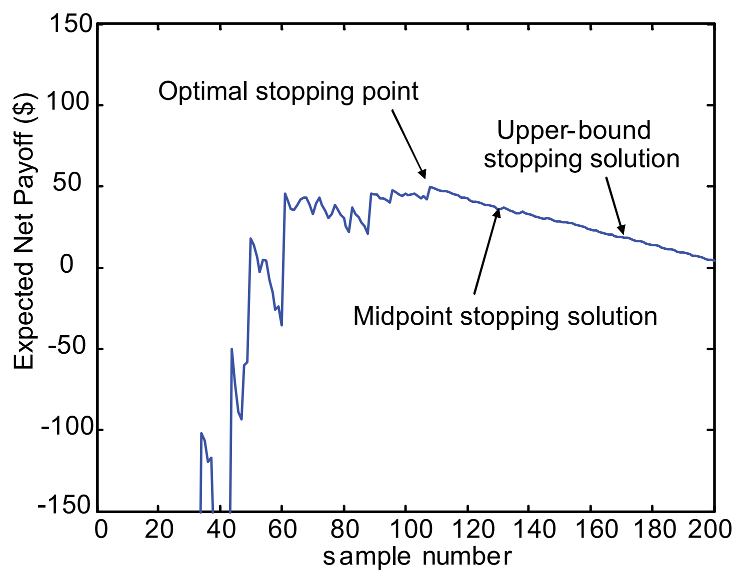


Figure 9.11. Actual expected net payoffs, additional trace

bounds, any policy that selects a point between the bounds for a single decision is acceptable, because the true value is only known to be somewhere between them.

For certain problems, a particular policy may tend to work better. For example, the Γ-maximax and midpoint policies are compared loosely for the pressure vessel design example, and it was found that the midpoint policy almost always performs better. If such results could be generalized to specific sets of problems, then designers might be able to choose an appropriate decision policy based on meta-information about the design problem. This would greatly increase the impact of the approach for bounding the value of information presented in this paper. Whether such generalizations are possible is an open research question.

Design decision policies. As explained in the 9.6.1, a design decision policy based on precise probabilities (the best-fit normal distribution) is chosen in order to focus on the effectiveness of using imprecise probabilities to represent the DM's state of information when estimating the value of information. Previous work has shown the value of incorporating the imprecision of the DM's state of information directly into the design decision policy (see Chapter 5). The use of imprecise probabilities for both the design decision policy and the prediction of the value of information appears to be a more realistic representation of a typical design problem and could possibly lead to additional insight.

Unknown Distributions. The p-box can be used to represent the DM's lack of knowledge about a distribution when the DM has varying amounts of initial knowledge about the distribution (see Section 4.3). In this paper, it was assumed that the DM knows the type of distribution but has no knowledge about the values of the distribution's parameters. It could be that the DM can only specify a set of all possible distribution types, or even that the DM lacks any knowledge about the possible distribution types. In both cases, the p-box would be significantly broader resulting in wider value bounds. The applicability of the new approach in other classes of problems could be explored by investigating how it performs under varying amounts of initial knowledge about the distribution being characterized.

Computational cost. The approach presented in this paper requires a double-loop Monte Carlo simulation for every sample size. For this experiment, the calculation of bounds on the value of the next sample takes about 5 minutes with a high number of p-box and message samples, though results for runs as short as 30 seconds appear nearly as good. These times are on a single 2.6 GHz Pentium 4 processor system with 512 MB of RAM.

Although this computation time seems perfectly reasonable, the computational complexity can be expected to increase substantially for more complicated design problems; hence, the proposed approach will need to be modified for application to complex design problems. For example, some p-box computations can be performed using algorithms with foundations in interval analysis that do not require second order Monte Carlo techniques (the DBC method described in Section 4.5.1) and are consequently much less computationally expensive on average (Ferson and Ginzburg 1996). Future work investigating how to adapt these methods for computing and simulating directly with p-boxes such that they can be used in the proposed approach is needed.

Design problems. The pressure vessel example used in this paper is deliberately simple, in order to illustrate concisely the information economic framework for information collection in design. To assess the general applicability of our approach, several variations of the design problem should be investigated. First, the material choice and hence the true distribution being characterized could be varied. Second, the payoff function could be varied to consider different levels of risk-preference. Third, a design problem in which there are multiple uncertain parameters and multiple sources of information could be explored. Modifications in the computational methods, as mentioned in the computational cost section, as well as possible modifications to the approach itself may be required in order to apply this approach in a computationally feasible manner to complex engineering design problems. These three design problem variations could lead to more general conclusions about the applicability of the approach.

9.10 Summary

In this chapter, the principles of information economics have been introduced and related to engineering design problems in which statistical parameters describing distributions are not fully known. The approach for applying information economics has been presented, several example scenarios and decision policies have been explored, limitations have been described, and areas for future work have been identified. An open question is how to make a decision given these bounds on value.

The main contribution of this work is the development of an approach by which the bounds on the value of information can be calculated by a designer during the information collection process using imprecise probabilities. This contribution can have a significant impact on engineering design by opening more problem classes to formal cost-benefit analysis during the problem formulation phase and information collection tasks of design decisions. Specifically, a formal cost-benefit analysis could help guide expenditures for information gathering in high-risk designs where difficult-to-characterize, uncommon events with severe consequences play a significant role in decisions.

CHAPTER 10:

DISCUSSION AND REMARKS

This chapter is a discussion and synthesis of material from the preceding chapters. The first goal is to reflect upon the material presented earlier by restating the motivating questions and answers proposed in Section 1.5 and by highlighting the main contributions. The second goal is to look outward and to the future. Throughout the dissertation, remarks were made with regard to future work. In this chapter, the limitations of the presented methods and needs for future work are reviewed from a broader perspective of practical use in engineering problems.

10.1 Review of motivating questions

The primary motivating question for the dissertation is:

How should engineering designers manage information to support decision making under uncertainty?

The basic answer presented in this dissertation is:

It is often valuable for engineers to represent the uncertainty in their information using probabilities that are most generally subjective and imprecise and to manage information collection and uncertainty modeling decisions using the principles of information economics.

This answer was developed in several steps, which can be broken into three parts. The first part, which is composed of Chapter 2 through Chapter 7 and forms the core of the dissertation, considers the nature of uncertainty and the value of using particular uncertainty models in engineering design. The second part, Chapter 8, addresses decision making in engineering design using the selected model. In the third part, Chapter 9, management of information collection is addressed. Each of these sections of the dissertation is reviewed in he following sections of this chapter. First, the contributions are reviewed in terms of the motivating questions (Sections 10.1.1 through 10.1.3). Then the main contributions of the dissertation are summarized in Section 10.2.

10.1.1 Contributions relating to modeling uncertainty

Effective decision making under uncertainty requires the DM to recognize the limits of the available information by modeling the uncertainty in the decision. The selection of an uncertainty model involves three important sub-questions:

1. What is the nature of uncertainty?

2. What is the most appropriate model of this uncertainty?

3. How should uncertainty models be compared?

These questions roughly correspond to Motivating Questions 1-3. The answers to these questions and the related conclusions are summarized in the following.

Motivating Question 1: What are the fundamental characteristics of uncertainty in the context of engineering design?

Answer 1: Uncertainty can have characteristics of being irreducible (or random) in nature as well as characteristics that are related to a imprecision in available knowledge—a reducible lack of knowledge.

Main conclusions: In Chapter 2, both the philosophical and practical literature on uncertainty was reviewed. Building from this literature and using simple examples, an argument was constructed that not all uncertainty that engineers face is random or irreducible in nature, and engineers therefore should recognize two characteristics of uncertainty: imprecision and irreducible uncertainty. Although this distinction is not always clear philosophically, the distinction is useful from a modeling perspective. The practical value of a method that represents imprecision distinctly from irreducible uncertainty in engineering design was demonstrated in Chapter 5. This practical comparison moves beyond the theoretical discussion and speculation in the existing literature.

Motivating Question 2: What is the most appropriate and general model of uncertainty for engineering design?

Answer 2: It is often valuable to use probabilities that are most generally subjective and imprecise.

Main conclusions:

- In Chapter 3, several uncertainty models that have been proposed in the literature were considered with particular emphasis on the existence of an operational definition for them and their ability to represent both irreducible uncertainty and imprecision. Imprecise probability theory is the only model with both a clear operational definition and the ability to represent both irreducible uncertainty and imprecision distinctly from each other.

- In Chapter 3, it was determined that if probability theory is to be used in engineering design, it is necessary to adopt a subjective interpretation of probability.

- In Chapter 4, the existing work on a sub-set of imprecise probability theory, called probability bounds analysis, was described and related to decision making in engineering design.

- In Chapter 5, a computational method for comparing the practical value of two uncertainty models was developed for the scenario of modeling limited statistical data. Using this method, it was shown that when imprecision is high (i.e. when the DM has little information), the use of PBA can lead to better designs than a precise, best-fit probability approach.

- In Chapter 6, the theoretical generality of probability bounds analysis as a sensitivity analysis for robustness is derived.

- In Chapter 7, the value—benefits, costs, and limitations—of using probability bounds analysis is compared to a decision analysis approach with sensitivity analysis with an example problem.

Motivating Question 3: How should engineers compare alternative models of uncertainty?

Answer 3: Designers should compare both the theoretical and practical validity of different models.

Main conclusions:

- In Chapter 4, the theoretical ability of imprecise probability theory to model the important characteristics of uncertainty is discussed, and this ability is compared to other proposed models.

- In Chapter 5, a computational method for comparing the practical value of two uncertainty models is developed for the scenario of modeling limited statistical data. This method is used as part of the validation of Answers 1 and 2.

- In Chapter 6, the theoretical generality of probability bounds analysis as an approach to sensitivity analysis for decision robustness is derived.

- In Chapter 7, the value—benefits, costs, and limitations—of using probability bounds analysis is compared to a decision analysis approach with sensitivity analysis with an example problem.

10.1.2 Contributions related to decision making

Motivating Question 4: How should a designer make a decision in the case of seemingly indeterminate information?

Answer 4: There is no generally optimal way to resolve inherent indeterminacy, but set-based decision rules exist that guide the efficient use of available information.

Main Conclusions: In Chapter 8, the decision criteria of interval dominance, maximality, and E-admissibility are described and demonstrated. Interval dominance is clearly valid to use in engineering design, and maximality can lead to additional, rational eliminations of alternatives by incorporating information about shared uncertainties. In the context of engineering design, it is rare that all imprecision will be eliminated before a final decision is taken. Consequently, E-admissibility is not a good policy because it may eliminate robust designs. It was also noted that the possible need to resort to arbitrary choice suggests an important direction for future work.

10.1.3 Contributions related to managing information collection

Motivating Question 5: What are the fundamental principles for managing information collection in engineering design?

Answer 5: Designers should apply the principles of information economics and can use methods that bound the value of information collection.

Main conclusions: In Chapter 9, the principles of information economics were introduced and related to the engineering design problem. A method was developed for bounding the value of additional information collection in the form of additional statistical samples. The method creates bounds on the value of the next sample, but in general it is necessary to resort to an arbitrary decision policy for resolving indeterminacy in the analysis. Consequently, the general applicability of the current method as an actual decision approach rather than a simple guide is limited.

10.2 Summary of contributions

This dissertation makes three significant contributions to engineering design under uncertainty, as summarized in the following subsections.

10.2.1 Contribution 1: Value of imprecise probabilities

The first contribution summarized is the core contribution of the dissertation. In this dissertation, the value of using an uncertainty model that explicitly recognizes imprecision and irreducible uncertainty was established from both practical and theoretical perspectives, and important limitations were noted.

1. The practical value—measured in terms of quality of the final design—of using PBA was compared to the practical value of using a traditional, best-fit precise model of uncertainty in a specific high-risk design problem involving limited statistical data.
    a. It was shown that when imprecision is high, the PBA approach yields designs that perform significantly better on average.
    b. It was shown that when imprecision is low, the PBA approach performs slightly worse on average.
    c. The identification of which state—high or low imprecision—is non-trivial for a specific design problem. The development of heuristics for guiding such judgments is an area for future work.
2. Via mathematical argument and example problem, it was demonstrated that PBA generalizes a global, all-way sensitivity analysis (as defined in decision analysis) and has its own advantages and limitations.
    a. When PBA is implemented using DBC, the PBA approach will not yield a Type I error, meaning that if the decision is sensitive to the existing lack of information (i.e. imprecision), then the PBA approach will always detect that sensitivity.

b. PBA can result in a Type II error, meaning that it can lead to the conclusion that there is sensitivity when there is not sensitivity.

c. Even with existing meta-sensitivity analysis approaches, PBA is severely limited in its ability to identify and prioritize the sources of sensitivity in a problem. Therefore, PBA does not help the DM prioritize information collection to reduce the sensitivity.

10.2.2 Contribution 2: Decision making with imprecise information

The second major contribution is in a crucial area for applying imprecise probabilities in engineering design—decision making in the presence of imprecision.

1. Existing decision criteria for eliminating imprecisely characterized decision alternatives from consideration in a design process were demonstrated.
2. The conclusion that one of the previously proposed approaches, E-admissibility, is not a good policy because it can remove robust solutions from consideration was reached.
3. Brought together the notion of set-based design with the decision policies for imprecise information, related these approaches to engineering design, and demonstrated their application in specific problem. As such, the preliminary foundations for a formal set-based approach to design in which the focus shifts from selection of the best alternative to the elimination of the inferior alternatives were established.

10.2.3 Contribution 3: Information collection and information economics

The third major contribution of the thesis involves the application of information economic principles to engineering design. This principle was embodied in the focus on the practical value of different uncertainty models and also explored more extensively in Chapter 9.

1. The principles of information economics were introduced into the engineering design process. Existing information economic concepts were extended to a formalized statement of the design process.
2. A method that uses imprecise probabilities to establish bounds on the value of future statistical data sample collection was developed. Previous methods assumed perfectly known joint and conditional probabilities. The new method lifts these restrictions and is a significant step forward towards the goal of formalizing the information economic trade-offs that designers should consider. However, the method requires the use of arbitrary choice, and consequently the results may not be generally application in a rigorous manner.

10.3 Onward and outward

As noted in the previous section, this dissertation contains several major contributions in engineering design. However, there are still significant opportunities for future work in the areas of imprecise probabilities, information economics, and the general area of engineering design under uncertainty. This dissertation is best viewed as a starting point—a starting point that will hopefully evolve into a turning point in engineering design.

Current practice in engineering design focuses on representing uncertainty using precise probabilities. This practice may have evolved for several reasons.

1. Engineers are generally exposed to traditional probability and statistics in their undergraduate programs, so there is a sense of familiarity with precise probabilities.

2. Precise probability theory comes with a neat, well-established decision theory—von Neumann-Morgenstern utility theory. The maximization of expected utility fits relatively nicely into optimization theory as it provides a single, precise objective function.
3. In general, precise probabilities are simpler to compute with than imprecise probabilities, so there may have been practical limitations on what representations were previously valuable for engineers to use.

None of these reasons are justifications for ignoring other uncertainty models. They can even be viewed as motivations for researching other models. As computing power increases, computing costs decrease, thus shifting the cost-benefit trade-off balances for particular design methods and uncertainty models. If a model promises significant benefits, then research can be focused on reducing the associated costs. This dissertation serves as a stepping stone towards future research by demonstrating this promise.

Specific benefits of using imprecise probabilities to capture both imprecision and irreducible uncertainty were established in this dissertation. Given these results, a complete dismissal of imprecise probabilities in engineering design cannot be made without acquiring significant evidence to support that conclusion. However, there are limitations to existing approaches for computing with imprecise probabilities in complex problems and making decisions with imprecise information that must be considered.

10.3.1 Considerations for complex problems

Many of the computing issues were discussed in Chapter 4. It was noted that recent work (Trejo and Kreinovich 2001, Kreinovich and Ferson 2004, Bruns 2006, Bruns and Paredis 2006, Bruns, et al. 2006) has examined and developed methods for applying PBA to black-box engineering models. However, these methods still require the analysis model to be evaluated a large number of times. For most computer models, this cost is likely to be prohibitive for the immediate future unless new methods are developed. For example, the replacement of complex analysis models with response surface models (Box and Wilson 1951, Myers and Montgomery 1995) and kriging models (Sacks, et al. 1989) is an important area for future work. Until such approaches are examined, the applicability of PBA to large scale engineering problems or problems with complex analysis models is an open issue.

The focus in this dissertation was on problems involving the acquisition of observable data, such as statistical data samples. However, a large source of information in the design of complex systems is expert opinion. Imprecise probabilities can theoretically model the uncertainty in such expert opinions, but the practicality of eliciting p-boxes from experts is questionable. Among other things, experts are not trained in creating p-boxes, and analysts are not trained in eliciting imprecise probabilities. For statistical problems, a large number of experts have experience in the combined art and science of fitting distributions and statistical models. There is no such expertise for p-boxes, so despite the promise of representing imprecise expert beliefs, there is currently no practical way of guiding the creation of such p-boxes. The general elicitation and assessment procedures introduced by Walley (1991) need significant refinement and extension for application in engineering design.

In a complex system, there are generally many sources of uncertainty. While PBA is great tool for performing sensitivity analysis for decision robustness identification, it is currently limited in its ability to prioritize information collection. In short, it informs the DM that he or she has a problem, but it provides no guidance as to how to fix the problem. To have a large impact on decision practice, an approach must be developed that captures at least some of the advantages of both PBA-based sensitivity analysis and the types of sensitivity analysis currently performed in decision analysis. These two approaches have complementary properties in that PBA is great at sensitivity analysis for decision robustness identification while the traditional approach is well suited for sensitivity analysis for information prioritization.

10.3.2 Considerations with regard to decision policies

Because imprecise probabilities generally lead to intervals of expected utility, traditional decision theory cannot be applied directly to analyses involving imprecise information. The potential overlap of intervals can lead to indeterminacy, a situation in which it is not possible to determine which decision alternative is the most preferred. Strictly speaking, this *should* lead to indecision because the available information provides no means for making a decision; that is exactly the point of using imprecise probabilities—to accurately model the decision maker's information state.

In a context such as set-based design in which the initial goal is just to reduce the design space, such indecision is acceptable provided some alternatives can be eliminated at every step. Essentially, the decision is recast from a selection of a preferred alternative to an individual decision about whether or not to eliminate that particular alternative. This will often work well up to a point, but eventually a single design must be chosen. In some cases, the elimination process may converge to a single alternative, but this will likely be the exception rather than the rule.

10.3.2.1 The use of arbitrary decision policies

Arbitrary decision policies have been suggested for resolving indeterminacy. These policies often result in single decisions that appear reasonable. However, the application of arbitrary decision policies in a sequence of decisions has not been studied in detail. When a DM applies an arbitrary decision rule, he or she is essentially selecting one point in the interval as the truth and using this truth to make a decision. This point corresponds to specific assumptions about the probabilities, probabilities that fall between the decision maker's upper and lower previsions. In order to maintain assumptions that lead to the proof of the avoidance of sure loss for coherent imprecise probabilities that was discussed in Section 3.4.3, a decision maker cannot later make a decision using a point that corresponds to different probabilities; the DM must be consistent in his or her selection of probabilities.

Initial investigations suggest that the consistent application of arbitrary policies is non-trivial. Perhaps the very notion of "arbitrary" implies "inconsistent." The practical consequences of using arbitrary policies in engineering design are unclear. In the context of gambling, the inconsistent selection of probabilities can result in a Dutch Book, a sequence of bets in which the DM is guaranteed to lose (see Section 3.4.3). This situation requires multiple decisions to be made—a situation that often exists in engineering design. Ideally, a decision policy would be applicable across multiple decisions, so this problem of arbitrary choice requires additional investigation.

10.3.2.2 The impact of arbitrary decision policies on this dissertation

The possibility of arbitrary decision policies leading to inconsistent behaviors does not undermine the primary results of this dissertation. The possible inconsistency in the use of arbitrary decision policies has no effect on the contributions regarding PBA as a sensitivity analysis because those methods do not rely on any decision policies. The sensitivity of the decision is identified by the presence of overlapping intervals; there is no need to reduce the intervals to points in order to identify the sensitivity. Consequently, the validity of arbitrary decision rules is irrelevant in this area.

The example problem in Chapter 5 was based on the assumption of a Γ-maximin policy, which does not lead to inconsistent behavior, as explained briefly in the following. When using a Γ-maximin policy, a DM essentially assumes that the worst case scenario[18] will occur and chooses the action that gives the best result in that worst case scenario. If a subsequent decision is made, the DM again considers the worst case scenario and makes a decision that leads to the best result in that worst case. Since the DM is always considering (and mitigating) the worst case, no subsequent decision could make him or her worse off (on average) that the lower bound on the expected value for the chosen decision, because that lower bound already is the worst case. Consequently, a Dutch Book cannot be constructed.

Viewed in terms of gambling behavior, a Γ-maximin policy is the most pessimistic policy possible. A very pessimistic DM would never be induced into a sequence of bets that result in a sure loss, because the pessimistic DM assumes that he or she will almost always lose. To the pessimist, both buying and selling look like worse decisions that they probably area. In order for a Dutch Book to be constructed, a DM must be optimistic in that he or she sells a bet at a price lower than reasonable and buys a bet at a price higher than reasonable. A pessimist does the opposite, buying only at a lower price that reasonable and selling only at a higher price than reasonable.

[18] The notion of "worst case" is defined in terms of the imprecision, such as the lowest bound on the expected utility. It is not "worst case" in terms of the worst single outcome possible, which would depend on any stochastic process, too. For example, consider an event with two outcomes, good (G) and bad (B). The probabilities of these two outcomes are not known precisely. It is known only that P(B)= [0.01, 0.05] and P(G)=[0.95,0.99]. In this context, the worst case is when P(B)=0.05 (as large as possible) and P(G)=0.95 (as small as possible). Even in the worst case, the good outcome can occur, and is in fact highly probable to occur in this example. Nevertheless, P(B)=0.05 and P(G)=0.95 are the worst case scenarios in terms of imprecision.

10.3.3 Potential areas for future application of imprecise probabilities and information economics

The preceding two sub-sections examined future research aimed at removing some of the existing limitations of using imprecise probabilities in engineering design. There are also many reasons to be optimistic, as there are many applications for imprecise probabilities beyond those of simple design selection, design elimination, and bounding the value of future information collection. Several application areas are explored in the following sub-sections.

10.3.3.1 Extensions to reliability and risk-based design

Reliability analysis and risk-based design are two important areas in which the use of imprecise probabilities may be valuable. In both of these areas, uncertainty plays a key role. For example, reliability analysis studies the probability of failure. When information is scarce, the probability of failure can be difficult to estimate. Imprecise probability theory presents a means for establishing bounds on probabilities instead of point estimates. However, the introduction of imprecise probabilities into traditional reliability analysis raises new questions, such as the propagation of imprecise reliability estimates through the analysis of systems and subsystems. This might be particularly valuable in situations in which experiments are prohibitively expensive, such as space craft design, or situations in which rare events are important. In these cases, expert opinion and generalizations from prior experience may be the only sources of information. A possible future research question is *how can subjective judgments be represented using imprecise probabilities in reliability analysis*?

A related question is the use of imprecise probabilities in reliability-based design optimization (RBDO). In RBDO, the designer seeks to maximize some performance function subject to probabilistic constraints. Recent literature has examined the use of alternatives to precise probabilities in RBDO, such as possibility theory (Mourelatos and Zhou 2005b) and evidence theory (Mourelatos and Zhou 2005a). In this work, upper bounds on the possibility of failure of constraints are used in the optimization rather than point estimates. The goal is to ensure that the possibility of failure is below the critical level, the scarcity of available information. The application of imprecise probabilities in RBDO is an important direction for future work, especially given the apparent benefits seen with other bounding approaches that do not have the rigor of the operational definition and the solid axioms provided by imprecise probability theory.

A basic goal of risk-based design is to reduce risk while meeting the other overall goals of a system. The risk associated with an action is generally defined in terms of the magnitude of the consequences of the action and the likelihood of those consequences (Stamatelatos 2000). The estimation of the likelihood of particular consequences is tied directly to the probabilities of particular states of the world. When information is scarce, it can be difficult to estimate these probabilities. It therefore may be valuable to apply imprecise probabilities to risk-based design and to extend probabilistic risk assessments to include measures of imprecision. A related research question is *in which aspects of risk analysis can these methods quantify risks that are currently dealt with qualitatively?* An answer to this question could improve significantly the accuracy and rigor of risk-based design approaches.

10.3.3.2 Set-based design methods for exploration of novel and imprecisely defined concepts.

As discussed in Chapter 8, engineers traditionally start with a large set of concepts and quickly narrow that set down to a single alternative that is iteratively modified and refined during embodiment and detail design. Because preliminary concepts are incompletely defined, a designer usually cannot determine an optimal concept; such a choice would require information about the details of the design that are not available until the design process is completed.

As mentioned in Chapter 8, some researchers have noted the success of set-based approaches to design in the automotive industry (Ward, et al. 1995, Sobek and Ward 1996). In these approaches, several concepts are designed in parallel. One advantage of these approaches is that creative, untested designs can be considered while simultaneously developing variants of existing, tried and true designs. Something novel is by definition new, and hence no one has experience with it. This lack of knowledge, or imprecision, makes predicting the exact performance of a novel concept impossible.

In this dissertation, methods for representing and reasoning with imprecision were developed and demonstrated, but the question remains as to how to manage a design process such that imprecision in several designs is systematically and efficiently reduced until one design can be chosen as the superior. By fostering the exploration of creative concepts, such a formal and rigorous set-based design method that considers multiple concepts in parallel and actively manages imprecision and information could lead to revolutionary advances in engineering technology that would otherwise be dismissed as too risky.

10.3.3.3 Model validation and uncertainty propagation.

The use of computer models and simulations for engineering analysis is widespread, a fundamental question remained unanswered: *how can a decision-maker develop sufficient confidence that a model is appropriate for guiding a particular decision*? Additional questions arise when models are reused or connected, either directly or in the form of a federated simulation. For example: *How can actions and responses be coordinated in a federated simulation when the uncertainty in the acting model is significantly different from the uncertainty in the reacting model*? More generally: *What are the fundamental methods for propagating uncertainty in federated simulations*?

The use of imprecise probabilities to represent the uncertainty associated with modeling simplifications and assumptions appears to be relatively novel research area. Recent work has considered the characterization of models using interval-based context and interval-based inaccuracy representations (Malak 2005), as well as the selection of an appropriate model from a set of models whose accuracy is characterized as intervals (Ling 2006). Based on these works, there may be ways to use imprecise probabilities to extend the process of model validation and assessment, just as imprecise probabilities enable an extension of application of information economics in engineering design, as described in Chapter 9.

10.3.3.4 Sensor tasking and data fusion.

Networks of distributed sensors are becoming quite common and important, especially in military applications. For example, information from radar, infrared, and visual spectrum sensors on various platforms must be integrated to create a full picture of battlefield. The entire purpose of sensors is to provide information, thereby reducing uncertainty. The information extracted from sensor data frequently is used to make time-critical decisions, such as neutralizing an incoming torpedo or missile. It is therefore essential to use available resources efficiently. This leads to the important question: *how can sensors be efficiently tasked to reduce uncertainty and quickly support tactical decision making*? A corollary question is *how can uncertain data from a network of various sensors be aggregated and processed quickly*? Non-military applications include monitoring plant performance (such as a refinery, manufacturing unit, or nuclear power facility) and implementing digital control systems.

The questions presented in the preceding paragraph are very complementary to the questions of information management in engineering design discussed in Chapter 9. Consequently, there is reason to think that research advances in the two areas may be mutually beneficial.

10.4 Revisiting the journey

The previous sections were structured reviews of the motivating questions, contributions, and areas for future work. In this section, the overall story of the dissertation and the cohesion of the different components is revisited.

In Section 1.5, the important elements of the general story of the dissertation was summarized with the following five questions:

1. When constructing a model, it is first necessary to understand what one is trying to model, in this case asking *what are the fundamental characteristics of uncertainty in the context of engineering design?*
2. Once the nature of uncertainty is described, what is the most appropriate and general model of uncertainty for engineering design?
3. Given several potential models from which to choose, how should engineers compare alternative models of uncertainty?
4. Once a model is chosen, how can this model be used to support the decisions?
5. Finally, given an existing state of knowledge, a model of uncertainty associated with it, and an approach for solving the original decision problem, how should one decide whether to collect more information (proceeding with the problem formulation phase) or whether to proceed to problem solution phase?

In this section, the story of the dissertation is retold in a summarized form via loose analogy. The goal of the section is to capture the essence of the dissertation.

10.4.1 Modeling uncertainty and making decisions: using the right tool for the right task

When an engineer approaches a design problem, he or she usually brings certain tools along. Each tool has specific benefits, costs, and limitations. The goal of this dissertation is to examine the tools available to the engineering as he or she explores the design of a product, focusing specifically on tools for modeling uncertainty.

Most engineers have precise probability theory strapped tightly to their utility belts, and this is the first tool that they reach for when they need to model uncertainty. In this dissertation, the engineer is taken back to the toolbox and encouraged to reexamine the available options. Before being allowed to examine the tools, the engineer is directed back to the planning phase and challenged to answer, "What is the task that you really want to perform?" Once the task—modeling uncertainty—is reviewed (in Chapter 2), the engineer is allowed to open the toolbox and select a tool (in Chapter 3 and Chapter 4).

The first few tools that the engineer examines—fuzzy sets and possibility theory—are very different from his or her trusted sidekick of precise probability theory. Unfortunately, they do not come with instructions on how to use them—they lack an operational definition. Additionally, as an uncertainty tool, they do not seem to be applicable to many of the tasks that probability theory completes with ease. The engineer once again takes probability theory from its home in his or her utility belt and admires its qualities.

However, the engineer still longs for more. Occasionally applying probability in engineering design is a bit like inserting a screw into wallboard with a hammer; the screw enters the wallboard, but the result is messy and not as sturdy as one would hope. The engineer ponders aloud, "Isn't there some way I can improve this tool without sacrificing its basic functionality?"

This is what imprecise probability theory does. It generalizes probability theory, including precise probabilities as a special case. It is as if the engineer's hammer now has a set of fold-out screwdrivers built in; it can do everything a hammer can do, and then some. The engineer decides to choose a tool—PBA—that is slightly less sophisticated but much easier to use than the general tool—imprecise probabilities. However, as anyone who has every ordered from a gadget catalogue probably knows, not everything that sounds great works out quite so well in practice.

In this dissertation, specific advantages of imprecise probabilities were demonstrated, both in terms of practical value of the added "tool" of modeling imprecision that imprecise probability theory adds to precise probability theory (Chapter 5 and Chapter 7) and in terms of a form of analysis based on imprecise probabilities to perform an existing function—sensitivity analysis—just as well or better than other tools (Chapter 6 and Chapter 7).

10.4.2 Managing uncertainty: exploration

The management of information and uncertainty in engineering design is not just a practical matter of finding a good tool and using it. It is also a philosophically challenging problem. Decision making under uncertainty is journey into a dark cave of the unknown. An explorer can take several approaches to traveling through this cave. For example, the explorer can crawl forward blindly, bumping into walls and frequently backtracking, essentially using just his or her hands and feet as tools for environmental exploration. This explorer might stumble into a hidden treasure or an exit now and then, but more often he or she will fall into an unforgiving crevice or forever be lost in an unyielding labyrinth of darkness.

A slightly smarter but less adventurous explorer may bring a tool along to assist with the journey, such as a flashlight. The flashlight reveals imminent perils and can illuminate exits from large caverns. As such, a flashlight provides specific benefits to the explorer. However, it comes with costs; the explorer must carry the weight of the flashlight and batteries, and the use of a flashlight may awaken pesky bats or stir intimidating insects. The benefits of a flashlight are also limited. No matter how powerful the beam, it will not reach around the next bend. When the DM comes to a fork in the cave, he or she will have to pursue one, and often without being able to see the end until committing to one and exploring it.

An engineer is an explorer in the design process. No matter what tool he or she uses to model uncertainty, the uncertainty will still exist. A good uncertainty model allows the decision maker to assess exactly what information is available, just as a flashlight allows an explorer to identify tunnels leading from a cavern and to avoid dangerous pitfalls. However, when two tunnels lead from a cavern and the engineer cannot see the end of either, the available information leads to indecision, and some other approach must be used to make a choice (Chapter 8).

Often when faced with indecision, an engineer can perform experiments or hire experts in order to acquire more information about the alternatives of a decision (Chapter 9). This process of acquiring information is also a bit like peering down a tunnel. When an engineer chooses to acquire information, he or she does not know what information will be received. If he or she did, then there probably is no reason to collect that information because it is already known. For example, consider the explorer peering down two possible exits from a cavern. The explorer can collect information about the decision by exploring one of the tunnels. However, the explorer has no idea what he or she will find; if he or she knew that *a priori*, there would be no reason to explore the cave to start with! In the design context, the engineer is left with an approach (Chapter 9) that bounds the value of additional data sample collection. This method provides guidance, but no clear answers.

Continuing the analogy from above, these bounds are to the engineer what a quick glimpse around the next bend of a tunnel may be to the explorer. A quite glimpse may reveal a very promising take on the situation, or it may yield a very discouraging outlook. Either way, the explorer cannot be sure where the tunnel leads. The idea is that by accounting for such glimpses, better decisions can be made, even if only in a heuristic manner. Previously, the methods for characterizing this type of information were limited, so the approach presented in this dissertation is a significant step forward and a strong starting point for future research.

10.5 Summary

The *establishment of strong foundations for future research* is a concise description of the overall contribution of this dissertation. In this dissertation, a pervasive challenge in engineering design was examined: the management of uncertainty in design decisions. For a variety of reasons, standard practice is to model uncertainty using traditional, precise probability theory. The familiarity and success of probability-based methods gives most designers little motivation to seek other models of uncertainty. This dissertation is an effort to motivate designers to consider other options.

In this dissertation, it was argued that there is more to uncertainty than what is captured by precise probabilities. It was shown that it is sometimes valuable to explicitly recognize imprecision (i.e. a lack of knowledge) in addition to probabilistic aspects by using a subset of imprecise probability theory called *probability bounds analysis* (PBA). It was also shown that PBA generalizes sensitivity analysis and allows for a rigorous identification of lack of robustness to imprecision in a particular decision. Preliminary approaches to decision making in the presence of imprecise information and to set-based design were presented, and an approach for guiding information collection decisions was developed.

These contributions form a foundation. The methods used in the dissertation are several steps from being employed in industrial practice. However, they are necessary and well-positioned stepping stones along the path to widespread use. The location of some future steps has been suggested, but ideally additional researchers will chart their own paths from the starting point built in this dissertation and revolutionize the *management of uncertainty in engineering design using imprecise probabilities and principles of information economics*.

REFERENCES

(2002). "Light Duty Oil Filters (U.S.)". Last accessed May 16, 2006, from <http://www.filtercouncil.org/envinfo/newsletter/Feb2002.pdf>.

(2006). "The NIST Reference on Constants, Units, and Uncertainty: International System of Units: Historical Context of the SI: Unit of Length (Meter)", The National Institute of Standards and Technology. Last accessed from <http://physics.nist.gov/cuu/Units/meter.html>.

Allen, J. K., C. C. Seepersad and F. Mistree (2006). "Robust Design for Multiscale and Multidisciplinary Applications." *Journal of Mechanical Design*, 128(4), 832-43.

Anscombe, F. J. and R. J. Aumann (1963). "A Definition of Subjective Probability." *Annals of Mathematical Statistics*, 34, 199-205.

Antonsson, E. K. and K. N. Otto (1995). "Imprecision in Engineering Design." *Journal of Mechanical Design*, 117, 25-32.

Arrow, K. and L. Hurwicz (1972). "An Optimality Criterion for Decision-Making under Ignorance." In C. F. Carter and J. L. Ford (Eds.), *Uncertainty and Expectation in Economics: Essays in Honour of G. L. S. Shackle*. Oxford: Blackwell.

Aughenbaugh, J. M., S. Duncan, C. J. J. Paredis and B. A. Bras (2006). "A Comparison of Probability Bounds Analysis and Sensitivity Analysis in Environmentally Benign Design and Manufacture." ASME 2006 International Design Engineering Technical Conferences, Design Automation Conference, Philadelphia, PA, ASME, Paper No. DETC2006-99230.

Aughenbaugh, J. M., J. M. Ling and C. J. J. Paredis (2005). "Applying Information Economics and Imprecise Probabilities to Data Collection in Design." 2005 ASME International Mechanical Engineering Congress and Exposition, Orlando, FL, USA, November 5-11, 2005, Paper No. IMECE2005-81181.

Aughenbaugh, J. M. and C. J. J. Paredis (2004). "The Role and Limitations of Modeling and Simulation in Systems Design." 2004 ASME International Mechanical Engineering Congress and Exposition, Anaheim, CA, USA, November 13-19, 2004, Paper No. IMECE2004-5981.

Aughenbaugh, J. M. and C. J. J. Paredis (2005). "The Value of Using Imprecise Probabilities in Engineering Design." ASME 2005 International Design Engineering Technical Conferences, Long Beach, CA, USA, September 24-28, 2005, Paper No. DETC2005-85354.

Aughenbaugh, J. M. and C. J. J. Paredis (2006a). "The Value of Using Imprecise Probabilities in Engineering Design." *Journal of Mechanical Design*, 128(4), 969-79.

Aughenbaugh, J. M. and C. J. J. Paredis (2006b). "Why Are Intervals and Imprecision Important in Engineering Design?" Reliable Engineering Computing Workshop, Savannah, GA, USA.

Aumann, R. J. (1962). "Utility Theory without the Completeness Axiom." *Econometrica*, 30, 445-62.

Aumann, R. J. (1964). "Utility Theory without the Completeness Axiom: A Correction." *Econometrica*, 32, 210-2.

Ayyub, B. (2001). *Elicitation of Expert Opinions for Uncertainty and Risks*. New York: CRC Press.

Balci, O. (1995). "Principles and Techniques of Simulation Validation, Verification and Testing." Winter Simulation Conference, (C. Alexopoulos, W. R. Lilegdon and D. Goldsman, Eds.), 147-54.

Bayes, T. (1764). "An Essay Towards Solving a Problem in the Doctrine of Chances." *Philosophical Transactions of the Royal Society of London*, 53, 370-418.

Bedford, T. and R. M. Cooke (2001). *Probabilistic Risk Analysis: Foundations and Methods*. New York: Cambridge University Press.

Ben-Haim, Y. (2001). *Information Gap Decision Theory*. London: Academic Press.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. (2nd ed.). Springer.

Berleant, D. (1993). "Automatically Verified Reasoning with Both Intervals and Probability Density." *Interval Computations*, 2, 48-70.

Berleant, D. and C. Goodman-Strauss (1998). "Bounding the Results of Arithmetic Operations on Random Variables of Unknown Dependency Using Intervals." *Reliable Computing*, 4(2), 147-65.

Berleant, D., L. Xie and J. Zhang (2003). "Statool: A Tool for Distribution Envelope Determination (DEnv), an Interval-Based Algorithm for Arithmetic on Random Variables." *Reliable Computing*, 9(2), 91-108.

Berleant, D. and J. Zhang (2004). "Representation and Problem Solving with Distribution Envelope Determination (DEnv)." *Reliability Engineering & System Safety*, 85(1-3), 153-68.

Bernoulli, J. (1713). "Ars Conjectandi."

Björklund, A. E. (2002). "Survey of Approaches to Improve Reliability in LCA." *International Journal of Life Cycle Analysis*, 7(2), 64-72.

Blanchard, B. S. (2004). *Systems Engineering Management*. (Third ed.). Hoboken, NJ: John Wiley and Sons, Inc.

Borel, E. (1909). *Elements De La Theorie Des Probabilities*. Paris: Librairie Scientifiques A. Hermann & Fils.

Box, G. E. P. (1979). "Robustness in the Strategy of Scientific Model Building." In R. L. Launer and G. N. Wilkinson (Eds.), *Robustness in Statistics*. New York: Academic Press.

Box, G. E. P. and K. B. Wilson (1951). "On the Experimental Attainment of Optimum Conditions." *Journal of the Royal Statistical Society, Series B*, 13, 1-45.

Bradley, S. R. and A. M. Agogino (1994). "An Intelligent Real Time Design Methodology for Component Selection: An Approach to Managing Uncertainty." *Journal of Mechanical Design*, 116(4), 980-8.

Bridgeman, P. W. (1927). *Logic of Modern Physics*. New York: The Macmillan Company.

Bruns, M. (2006). *Propagation of Imprecise Probabilities through Black Box Models*. Masters Thesis, Georgia Institute of Technology.

Bruns, M. and C. J. J. Paredis (2006). "Numerical Methods for Propagating Imprecise Uncertainty." ASME 2006 International Design Engineering Technical Conferences and Computers in Information Engineering Conference, Philadelphia, PA, Sept. 10-13, 2006, Paper No. DETC2006-99237.

Bruns, M., C. J. J. Paredis and S. Ferson (2006). "Computational Methods for Decision Making Based on Imprecise Information." Reliable Engineering Computing Workshop, Savannah, GA, USA.

Buede, D. M. (2000). *The Engineering Design of Systems: Models and Methods*. New York: John Wiley & Sons, Inc.

Byrne, D. M. and S. Taguchi (1987). "The Taguchi Approach to Parameter Design." 40th Annual Quality Congress Transactions.

Cannon, C. M. and Z. W. Kmietowicz (1974). "Decision Theory and Incomplete Knowledge." *Journal of Management Studies*, 11(3), 224-32.

Carnahan, J. V., D. L. Thurston and T. Liu (1994). "Fuzzy Ratings for Multiattribute Design Decision-Making." *Journal of Mechanical Design*, 116(2), 511.

Carnap, R. (1950). *Logical Foundations of Probability*. Chicago: Chicago University Press.

Casti, J. L. (1990). *Searching for Certainty*. New York: William Morrow.

Cheeseman, P. (1988). "An Inquiry into Computer Understanding (with Responses)." *Computational Intelligence*, 4, 58-66.

Chen, W., J. K. Allen, D. N. Marvis and F. Mistree (1996). "A Concept Exploration Method for Determining Robust Top-Level Specifications." *Engineering Optimization*, 26, 137-58.

Chen, W., R. Jin and A. Sudjianto (2004). "Analytical Variance-Based Global Sensitivity Analysis in Simulation-Based Design under Uncertainty." ASME 2005 International Design Engineering Technical Conferences, Salt Lake City, UT, USA, September 28-October 2, 2004, Paper No. DETC2004-57484.

Ciocoiu, M., M. Gruninger and D. Nau (2001). "Ontologies for Integrating Engineering Applications." *ASME Journal of Computing and Information Science in Engineering*, 1(1), 12-22.

Clemen, R. T. (1996). *Making Hard Decisions: An Introduction to Decision Analysis*. (Second ed.). New York: Duxbury Press.

Cooke, R. (2004). "The Anatomy of the Squizzel - the Role of Operational Definitions in Representing Uncertainty." *Reliability Engineering & System Safety*, 85(1-3), 313-9.

Cooke, R. M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Oxford: Oxford University Press.

Coolen, E. P. A. (2004). "On the Use of Imprecise Probabilities in Reliability." *Quality and reliability in engineering international*, 20, 193-202.

Cooper, L. P. (2003). "A Research Agenda to Reduce Risk in New Product Development through Knowledge Management: A Practitioner Perspective." *Journal of Engineering and Technology Management*, 20, 117-40.

Cox, R. T. (1946). "Probability, Frequency and Reasonable Expectation." *American Journal of Physics*, 14, 1-13.

Cullen, A. C. and H. C. Frey (1999). *Probabilistic Techniques in Expose Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. New York: Plenum Press.

Dai, Z., M. J. Scott and Z. P. Mourelatos (2003). "Incorporating Epistemic Uncertainty in Robust Design." ASME 2003 International Design Engineering Technical Conferences, Design Automation Conference, Chicago, IL, United States, Sept. 2-6, 2003, American Society of Mechanical Engineers, Paper No. DETC2003/DAC-48713.

de Cooman, G. and P. Walley (2002). "A Possibilistic Hierarchical Model for Behaviour under Uncertainty." *Theory and Decision*, 52(4), 327.

de Finetti, B. (1974). *Theory of Probability Volume 1: A Critical Introductory Treatment*. (Vol. 1). New York: Wiley.

de Finetti, B. (1975). *Theory of Probability Volume 2*. (Vol. 2). Wiley.

de Finetti, B. (1980). "Foresight. Its Logical Laws, Its Subjective Sources (Translated)." In H. E. Kyburg and H. E. Smokler (Eds.), *Studies in Subjective Probability*. E. Krieger Publishing Company. (Original work published in La Prevision: Ses Lois Logiques, Ses Sources Subjectives. Annales de l'Institute Henri Poincare 7:1-68, 1937).

Dempster, A. P. (1967). "Upper and Lower Probabilities Induced by a Multivalued Mapping." *The Annals of Statistics*, 28, 325-39.

Der Kiureghian, A. (1989). "Measures of Structural Safety under Imperfect States of Knowledge." *ASCE Journal of Structural Engineering*, 115(5), 1119-39.

Devore, J. L. (1995). *Probability and Statistics for Engineering and the Sciences*. (Fourth ed.). New York: Duxbury Press.

Du, L. and K. K. Choi (2006). "A New Fuzzy Analysis Method for Possibility-Based Design Optimization." *AIAA Journal*, to appear

Dubois, D. and H. Prade (1988). *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York: Plenum Press.

Efron, B. and C. N. Morris (1973). "Stein's Estimation Rule and Its Competitors--an Empirical Bayes Approach." *Journal of the American Statistical Association*, 68, 117-30.

Efron, B. and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability. New York: Chapman & Hall.

Elishakoff, I. (2004). *Safety Factors and Reliability: Friends or Foes?* Dordrecht, Netherlands: Kluwer Academic Publishers.

Eschenbach, T. G. (1992). "Spiderplots Versus Tornado Diagrams for Sensitivity Analysis." *Interfaces*, 22(6), 40.

Felli, J. C. and G. B. Hazen (2004). "Javelin Diagrams: A Graphical Tool for Probabilistic Sensitivity Analysis." *Decision Analysis*, 1(2), 93-107.

Fernández, M. C., C. C. Seepersad, D. W. Rosen, J. K. Allen and F. Mistree (2001). "Utility-Based Decision Support for Selection in Engineering Design." ASME 2001 Design Engineering Technical Conferences, Design Automation Conference, Pittsburgh, PA, Paper No. DETC2001/DAC-21106.

Ferson, S. (2000). "Probability Bounds Analysis Solves the Problem of Incomplete Specification in Probabilistic Risk and Safety Assessments." Ninth Conference on Risk-Based Decision-making in Water Resources, Santa Barbara CA, USA.

Ferson, S. (2002). *RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Boca Raton, Florida: Lewis Publishers.

Ferson, S. and S. Donald (1998). "Probability Bounds Analysis." In A. Mosleh and R. A. Bari (Eds.), *Probabilistic Safety Assessment and Management* (pp. 1203-8). New York, NY: Springer-Verlag.

Ferson, S., L. Ginzburg, V. Kreinovich, H. T. Nguyen and S. A. Starks (2002a). "Uncertainty in Risk Analysis: Towards a General Second-Order Approach Combining Interval, Probabilistic, and Fuzzy Techniques." 2002 IEEE International Conference on Fuzzy Systems: FUZZ-IEEE'02, May 12-17 2002, Honolulu, HI, United States, Institute of Electrical and Electronics Engineers Inc., 1342-7.

Ferson, S. and L. R. Ginzburg (1996). "Different Methods Are Needed to Propagate Ignorance and Variability." *Reliability Engineering & System Safety*, 54(2-3), 133-44.

Ferson, S., J. Hajagos, D. S. Myers and W. T. Tucker (2005). "Constructor: Synthesizing Information About Uncertain Variables." SAND2005-3769. Sandia National Laboratories, Albuquerque, NM.

Ferson, S., J. Hajagos and W. T. Tucker (2004a). "Probability Bounds Analysis Is a Global Sensitivity Analysis." International Conference on Sensitivity Analysis of Model Output (SAMO), Sante Fe, New Mexico, March 8-11, 2004.

Ferson, S. and J. G. Hajagos (2004). "Arithmetic with Uncertain Numbers: Rigorous and (Often) Best Possible Answers." *Reliability Engineering & System Safety*, 85(1-3), 135.

Ferson, S., V. Kreinovich, L. Ginzburg, D. S. Myers and K. Sentz (2002b). "Constructing Probability Boxes and Dempster-Shafer Structures." SAND2002-4015. Sandia National Laboratories, Albuquerque, New Mexico.

Ferson, S., R. B. Nelsen, J. Hajagos, D. Berleant, J. Zhang, W. T. Tucker, L. Ginzburg and W. L. Oberkampf (2004b). "Dependence in Probabilistic Modeling, Dempster-Shafer Theory, and Probability Bounds Analysis." SAND2004-3072. Sandia National Laboratories, Albuquerque.

Ferson, S. and W. T. Tucker (2006 (in preparation)). "Sensitivity in Risk Analyses with Uncertain Numbers." Sandia National Laboratories, Albuquerque, NM. <http://www.ramas.com/sensanal.pdf>.

Finch, W. W. (1997). "A Set-Based System for Eliminating Infeasible Designs in Engineering Problems Dominated by Uncertainty." ASME Design Engineering Technical Conferences, Sacramento, California, September 14-17, ASME, Paper No. DETC97/DTM-3886.

Fishburn, P. (1964). *Decision and Value Theory*. New York: J. Wiley.

Fishburn, P. (1982). *Foundations of Expected Utility*. Dordrecht: D. Reidel.

Forsberg, K. and H. Mooz (1992). "The Relationship of Systems Engineering to the Project Cycle." *Engineering Management Journal*, 4(3), 36-43.

Forsberg, K., H. Mooz and H. Cotterman (2000). *Visualizing Project Management: A Model for Business and Technical Success*. New York: Wiley.

Frey, D. D., G. Reber and Y. Lin (2005). "A Quadrature-Based Sampling Technique for Robust Design with Computer Models." ASME 2005 International Design Engineering Technical Conferences, Long Beach, CA, USA, September 24-28, 2005, Paper No. DETC2005-85490.

Gärdenfors, P. and N.-E. Sahlin (1982). "Unreliable Probabilities, Risk Taking, and Decision Making." *Synthese*, 53, 361-86.

Gärdenfors, P. and N.-E. Sahlin (1983). "Decision Making with Unreliable Probabilities." *British Journal of Mathematical & Statistical Psychology*, 36, 240-51.

Garvey, P. R. (1999). *Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective*. New York: Marcel Dekker, Inc.

Geer, J. F. and G. J. Klir (1992). "A Mathematical Analysis of Information-Preserving Transformations between Probabilistic and Possibilistic Formulations of Uncertainty." *International Journal of General Systems*, 20(2), 143.

Gelman, A., J. B. Carlin, H. S. Stern and D. B. Rubin (1995). *Bayesian Data Analysis*. London: Chapman and Hall.

Giles, R. (1982). "Foundations for a Theory of Possibility." In M. M. Gupta and E. Sanchez (Eds.), *Fuzzy Information and Decision Processes* (pp. 183-95). Amsterdam: North-Holland Publishing Company.

Goedkoop, M. and R. Spriensma (2001). "The Eco-Indicator 99: A Damage Oriented Method for Life Cycle Impact Assessment." PRé Consultants, Amersfoort, Netherlands. <http://www.pre.nl/eco-indicator99/ei99-reports.htm>.

Good, I. J. (1980). "Some History of the Hierarchical Bayesian Methodology." In J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith (Eds.), *Bayesian Statistics* (pp. 489-519) (Vol. 1). Valencia: Valencia University Press.

Good, I. J. (1983). *Good Thinking: The Foundations of Probability and Its Applications*. Minneapolis: University of Minnesota Press.

Groen, F. J. and A. Mosleh (2005). "Foundations of Probabilistic Inference with Uncertain Evidence." *International Journal of Approximate Reasoning*, 39(1), 49-83.

Gupta, S. K. and C. Xu (2002). "Estimating the Optimal Number of Alternatives to Be Explored in Large Design Spaces: A Step Towards Incorporating Decision Making Cost in Design Decision Models." ASME 2002 DETC CIE, Montreal, Canada, Paper No. DETC2002/CIE-34491.

Hajek, A. (2003). "Interpretations of Probability". Last accessed 2004, from <http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=probability-interpret>.

Hanoch, G. and H. Levy (1969). "The Efficiency Analysis of Choices Involving Risk." *Review of economic studies*, 36, 335-46.

Hardcastle, G. L. (1995). "S. S. Stevens and the Origins of Operationism." *Philosophy of Science*, 62(3), 404-24.

Hart, A. G. (1942). "Risk, Uncertainty and the Unprofitability of Compounding Probabilities." In O. Lange, F. McIntyre and T. O. Yntema (Eds.), *Studies in Mathematical Economics and Econometrics*. Chicago: University of Chicago Press.

Haukaas, T. (2003). "Types of Uncertainty, Elementary Data Analysis, Set Theory." Reliability and Structural Safety: Lecture Notes University of British Columbia, Vancouver, B.C.

Hazelrigg, G. (2003). "Thoughts on Model Validation for Engineering Design." ASME 2003 DETC, Chicago Illinois.

Hazelrigg, G. A. (1996). *Systems Engineering: An Approach to Information-Based Design*. Englewood Cliffs, NJ: Prentice Hall.

Hazelrigg, G. A. (1998). "A Framework for Decision-Based Design." *Journal of Mechanical Design*, 120(4), 653-8.

Helton, J. C. (1994). "Treatment of Uncertainty in Performance Assessments for Complex Systems." *Risk Analysis*, 14, 483-511.

Helton, J. C. and F. J. Davis (2000). "Sampling-Based Methods for Uncertainty and Sensitivity Analysis." SAND99-2240. Sandia National Laboratories, Albuquerque, New Mexico.

Hines, W. W., D. C. Montgomery, D. M. Goldsman and C. M. Borror (2003). *Probability and Statistics in Engineering*. (Fourth Edition ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Hofer, E. (1996). "When to Separate Uncertainties and When Not to Separate." *Reliability Engineering & System Safety*, 54(2-3), 113-8.

Hoffman, F. O. and J. S. Hammonds (1994). "Propagation of Uncertainty in Risk Assessments: The Need to Distinguish between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability." *Risk Analysis*, 14(5), 707-12.

Hofstetter, P. (1998). *Perspectives in Life Cycle Impact Assessment; a Structured Approach to Combine Models of the Technosphere, Ecosphere, and Valuesphere*. Boston: Kluwers Academic Publishers.

Homma, T. and A. Saltelli (1996). "Importance Measures in Global Sensitivity Analysis of Nonlinear Models." *Reliability Engineering & System Safety*, 52, 1-17.

Hora, S. C. (1996). "Aleatory and Epistemic Uncertainty in Probability Elicitation with an Example from Hazardous Waste Management." *Reliability Engineering & System Safety*, 54(2-3), 217-23.

Howard, R. A. (1965). "Bayesian Decision Models for Systems Engineering." *IEEE Transactions on Systems Science and Cybernetics*, SSC-1(1), 36-40.

Howard, R. A. (1966). "Information Value Theory." *IEEE Transactions on Systems Science and Cybernetics*, SEC-2(1), 22.

Howard, R. A. (1968). "Foundations of Decision Analysis." *IEEE Transactions on Systems Science and Cybernetics*, SSC-4(3), 211.

Howard, R. A. (1988a). "Decision Analysis: Practice and Promise." *Management Science*, 34(6), 679-95.

Howard, R. A. (1988b). "The Evolution of Decision Analysis." In R. A. Howard and J. E. Matheson (Eds.), *Readings on the Principles and Applications of Decision Analysis* (pp. 1-16) (Vol. I). Strategic Decisions Group. (Original work published in 1983).

Huber, P. J. (1981). *Robust Statistics*. New York: Wiley.

Jayaram, S., J. Vance, R. Gadh, J. Jayaraman and H. Srinivasan (2001). "Assessment of VR Technology and Its Applications to Engineering Problems." *ASME Journal of Computing and Information Science in Engineering*, 1(1)

Jeffreys (1939). *Theory of Probability*. Oxford Press.

Joslyn, C. A. and J. M. Booker (2005). "Generalized Information Theory for Engineering Modeling and Simulation." In E. Nikolaidis, D. M. Ghiocel and S. Singhal (Eds.), *Engineering Design Reliability Handbook* (pp. 9.1-9.40). New York: CRC Press.

Kacprzyk, J. and M. Fedrizzi (1990). *Multiperson Decision Making Models Using Fuzzy Sets and Possibility Theory*. Boston, MA: Kluwer Academic Publishers.

Kahneman, D., P. Slovic and A. Tversky (Eds.) (1982). *Judgment under Uncertainty. Heuristics and Biases.* Cambridge, UK.: Cambridge University Press.

Kearfott, R. B. and V. Kreinovich (1996). *Applications of Interval Computations*. (Vol. 3). AH Dordrecht, The Netherlands: Kluwer Academic Publishers.

Keeney, R. L. and H. Raiffa (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York, NY: Cambridge University Press.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan.

Kirkwood, C. W. and R. K. Sarin (1985). "Ranking with Partial Information: A Method and an Application." *Operations Research*, 33, 38-48.

Klir, G. and G. Yuan (1995). *Fuzzy Sets and Fuzzy Logic*. Upper Saddle River, NJ: Prentice Hall.

Klir, G. J. (1992). "Probabilistic Versus Possibilistic Conceptualization of Uncertainty." In B. M. Ayyub, M. M. Gupta and L. N. Kanal (Eds.), *Analysis and Management of Uncertainty: Theory and Applications*. New York: North-Holland.

Klir, G. J. and R. M. Smith (2001). "On Measuring Uncertainty and Uncertainty-Based Information: Recent Developments." *Annals of Mathematics and Artificial Intelligence*, 32(1-4), 5-33.

Kmietowicz, Z. W. and A. D. Pearman (1981). *Decision Theory and Incomplete Knowledge*. Aldershot, England: Gower.

Kmietowicz, Z. W. and A. D. Pearman (1984). "Decision Theory, Linear Partial Information and Statistical Dominance." *Omega*, 12, 391-9.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Boston: Houghton Mifflin.

Kolmogorov, A. N. (1956). *Foundations of Probability*. (N. Morrison. Trans.). New York: Chelsea Publishing Company. (Original work published in 1933).

Kreinovich, V. and S. Ferson (2004). "A New Cauchy-Based Black-Box Technique for Uncertainty in Risk Analysis." *Reliability Engineering & System Safety*, 85, 267-79.

Kyburg, H. E. (1987). "Objective Probabilities." IJCAI-87, 902-4.

Kyburg, H. E., Jr. and M. Pittarelli (1996). "Set-Based Bayesianism." *IEEE Transactions on Systems, Man & Cybernetics, Part A (Systems & Humans)*, 26(3), 324.

Laplace, P. S. (1951). *A Philosophical Essay on Probabilities*. New York: Dover Publications. (Original work published in 1814).

Law, A. M. and D. W. Kelton (2000). *Simulation Modeling and Analysis*. (Third ed.) McGraw-Hill Series in Industrial Engineering and Management Science. St. Louis: McGraw-Hill.

Lawrence, D. B. (1999). *The Economic Value of Information*. New York: Springer-Verlag.

Layard, R. and S. Glaister (Eds.) (1994). *Cost-Benefit Analysis (2nd Edition)*. Cambridge: Cambridge University Press.

Leamer, E. E. (1990). "Sensitivity Analysis Would Help." In C. W. J. Granger (Ed.), *Modelling Economic Series*. Oxford: Clarendon Press.

Lee, C. C. (1990). "Fuzzy Logic in Control Systems: Fuzzy Logic Controller." *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2), 419-35.

Levi, I. (1974). "On Indeterminate Probabilities." *Journal of Philosophy*, 71, 391-418.

Levi, I. (1980). *The Enterprise of Knowledge, an Essay on Knowledge, Credal Probability, and Chance*. Cambridge, MA: MIT Press.

Levy, H. (1998). *Stochastic Dominance: Investment Decision Making under Uncertainty*. Studies in Risk and Uncertainty. Boston: Kluwer Academic Publishers.

Li, X. and D. D. Frey (2005). "A Study of Factor Effects in Data from Factorial Experiments." ASME 2005 International Design Engineering Technical Conferences, Long Beach, CA, USA, September 24-28, 2005, Paper No. DETC2005-85486.

Lieberman, M. and R. Hall (2000). *Introduction to Economics*. South-Western College Publishing.

Lindley, D. V. (1982a). "Scoring Rules and the Inevitability of Probability." *International Statistical Review*, 50, 1-16.

Lindley, D. V. (1982b). "Subjectivist View of Decision Making." *European Journal of Operational Research*, 9(3), 213.

Lindley, D. V. (2000). "The Philosophy of Statistics." *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(3), 293.

Ling, J. M. (2006). *Managing Information Collection in Simulation-Based Design*. Master of Science in Mechanical Engineering Thesis, Georgia Institute of Technology.

Ling, J. M., J. M. Aughenbaugh and C. J. J. Paredis (2006). "Managing the Collection of Information under Uncertainty Using Information Economics." *Journal of Mechanical Design*, 128(4), 980-90.

Liu, H. and W. Chen (2006). "Relative Entropy Based Method for Probabilistic Sensitivity Analysis in Engineering Design." *Journal of Mechanical Design*, in press

Lu, B. and P. Gu (2003). "Systematic Life Cycle Design for Sustainable Product Development." ASME 2003 International Design Engineering Technical Conferences, Chicago, Illinois, USA, September 2-6, 2003, Paper No. DETC2003/DFM-48141.

Malak, R. J., Jr. (2005). *A Framework for Validating Reusable Behavioral Models in Engineering Design*. Masters Thesis, Georgia Institute of Technology.

Malak, R. J., Jr. and C. J. J. Paredis (2004). "On Characterizing and Assessing the Validity of Behavior Models and Their Predictions." ASME 2004 DETC, Salt Lake City, DETC2004-57452.

Marschak, J. (1974). *Economic Information, Decision, and Prediction: Selected Essays*. (Vol. I-III). Boston: D. Reidel Publishing Company.

Marston, M., J. K. Allen and F. Mistree (2000). "The Decision Support Problem Technique: Integrating Descriptive and Normative Approaches in Decision Based Design." *Engineering Valuation and Cost Analysis*, 3, 107-29.

Matheson, J. E. (1968). "The Economic Value of Analysis and Computation." *IEEE Transactions on Systems Science and Cybernetics*, SSC-4(3), 325-32.

McGovern, J., D. Samson and A. Wirth (1993). "Influence Diagrams for Decision Analysis." In S. Nagel (Ed.), *Computer-Aided Decision Analysis* (pp. 107-22). Westport, Connecticut: Quorum Books.

Merriam-Webster (1993). *Merriam-Webster's Collegiate Dictionary*. (10th Edition ed.). Springfield, MA: Merriam-Webster.

Mistree, F., W. F. Smith, B. A. Bras, J. K. Allen and D. Muster (1990). "Decision-Based Design. A Contemporary Paradigm for Ship Design." 1990 SNAME Annual Meeting, Oct 31-Nov 3 1990, San Francisco, CA, USA, Soc of Naval Architects & Marine Engineers, Jersey City, NJ, USA, 565-97.

Moore, R. E. (1966). *Interval Analysis*. Prentice-Hall Series in Automatic Computation. Englewood Cliffs, NJ: Prentice-Hall.

Moore, R. E. (1979). *Methods and Applications of Interval Analysis*. Philadelphia: Society for Industrial and Applied Mathematics.

Mourelatos, Z. and J. Zhou (2005a). "A Design Optimization Method Using Evidence Theory." ASME 2005 International Design Engineering Technical Conferences, Long Beach, CA, September 24-28, 2005, Paper No. DETC2005-84693.

Mourelatos, Z. P. and J. Liang (2004). "An Efficient Unified Approach for Reliability and Robustness in Engineering Design." NSF Workshop on Reliable Engineering Computing, Savannah, Georgia, September 15-17, 2004.

Mourelatos, Z. P. and J. Zhou (2005b). "Reliability Estimation and Design with Insufficient Data Based on Possibility Theory." *AIAA Journal*, 43(8), 1696.

Muhanna, R. L. and R. L. Mullen (2004). "Interval Methods for Reliable Computing." In E. Nikolaidis, D. M. Ghiocel and S. Singhal (Eds.), *Engineering Design Reliability Handbook*. New York, NY: CRC Press.

Murphy, T. E., K.-L. Tsui and J. K. Allen (2005). "A Review of Robust Design Methods for Multiple Responses." *Research in Engineering Design*, 15(4), 201.

Myers, R. H. and D. C. Montgomery (1995). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. New York: John Wiley and Sons.

Nau, R. F. (1992). "Indeterminate Probabilities on Finite Sets." *The Annals of Statistics*, 20, 1737-67.

Neter, J., M. H. Kutner, C. J. Nachtscheim and W. Wasserman (1996). *Applied Linear Statistical Models*. (4 ed.). St. Louis: WCB McGraw-Hill.

Nikolaidis, E. (2005). "Types of Uncertainty in Design Decision Making." In E. Nikolaidis, D. M. Ghiocel and S. Singhal (Eds.), *Engineering Design Reliability Handbook*. New York: CRC Press.

Nikolaidis, E., S. Chen, H. Cudney, R. T. Haftka and R. Rosca (2004). "Comparison of Probability and Possibility for Design against Catastrophic Failure under Uncertainty." *Journal of Mechanical Design*, 126(3), 386-94.

Oberkampf, W., J. Helton, S. Wojtkiewicz, C. Joslyn and S. Ferson (2002a). "Epistemic Uncertainty Workshop." *Reliable Computing*, 8(6), 503-5.

Oberkampf, W. L., S. M. DeLand, B. M. Rutherford, K. V. Diegert and K. F. Alvin (2002b). "Error and Uncertainty in Modeling and Simulation." *Reliability Engineering and System Safety*, 75(3), 333-57.

Oberkampf, W. L. and J. C. Helton (2002). "Investigation of Evidence Theory for Engineering Applications." Non-Deterministic Approaches Forum, 43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Denver, CO, April 21-23, 2002.

Oberkampf, W. L., J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz and S. Ferson (2001). "Challenge Problems: Uncertainty in System Response Given Uncertain Parameters." <http://www.sandia.gov/epistemic/ >.

Otto, K. N. and E. K. Antonsson (1991). "Trade-Off Strategies in Engineering Design." *Research in Engineering Design*, 3(2), 87.

Otto, K. N. and E. K. Antonsson (1992). "The Method of Imprecision Compared to Utility Theory for Design Selection Problems." *ASME 1993 Design Theory and Methodology Conference*

Pahl, G. and W. Beitz (1996). *Engineering Design: A Systematic Approach*. London: Springer Publishing.

Pan, Y. and B. Yuan (1997). "Bayesian Inference of Fuzzy Probabilities." *International Journal of General Systems*, 26, 73-90.

Parry, G. W. (1996). "The Characterization of Uncertainty in Probabilistic Risk Assessment of Complex Systems." *Reliability Engineering and System Safety*, 54(2-3), 119-26.

Parunak, H. V. D., A. Ward, M. Fleisher and J. Sauter (1997). "A Marketplace of Design Agents for Distributed Concurrent Set-Based Design." ISPE International Conference on Concurrent Engineering: Research and Applications (ISPE/CE97), Troy, Michigan, August 20-22, 1997, ISPE.

Pate-Cornell, M. E. (1996). "Uncertainties in Risk Analysis: Six Levels of Treatment." *Reliability Engineering & System Safety*, 54(2-3), 95-111.

Pedersen, K., J. Emblemsvag, R. Bailey, J. K. Allen and F. Mistree (2000). "Validating Design Methods and Research: The Validation Square." AMSE DETC 2000, Baltimore, MD.

Popper, K. R. (1957). "The Propensity Interpretation of the Calculus of Probability and the Quantum Theory." *The Colston Papers*, 9, 65-70.

Popper, K. R. (1959a). *The Logic of Scientific Discovery*. New York: Basic Books.

Popper, K. R. (1959b). "The Propensity Interpretation of Probability." *British journal of the philosophy of science*, 10, 25-42.

Pratt, J. W., H. Raiffa and R. Schlaifer (1995). *Introduction to Statistical Decision Theory*. Cambridge, MA: The MIT Press.

Radhakrishnan, R. and D. A. McAdams (2005). "A Methodology for Model Selection in Engineering Design." *Journal of Mechanical Design*, 127(May), 378-87.

Raiffa, H. (1968). *Decision Analysis, Introductory Lectures on Choices under Uncertainty*. Menlo Park: Addison-Wesley.

Ramsey, F. P. (1926). "Truth and Probability." In R. B. Braithwaite (Ed.), *The Foundations of Mathematics and Other Logical Essays*. London: Kegan Paul.

Rangan, R. M. and B. Chadha (2001). "Engineering Information Management to Support Enterprise Business Processes." *ASME Journal of Computing and Information Science in Engineering*, 1(1), 32-40.

Regan, H. M., S. Ferson and D. Berleant (2004). "Equivalence of Methods for Uncertainty Propagation of Real-Valued Random Variables." *International Journal of Approximate Reasoning*, 36(1), 1-30.

Reichenbach, H. (1949). *The Theory of Probability*. Berkeley: University of California Press.

Rekuc, S. J. (2005). *Eliminating Design Alternatives under Interval-Based Uncertainty*. Masters Thesis, Georgia Institute of Technology.

Rekuc, S. J., J. M. Aughenbaugh, M. Bruns and C. J. J. Paredis (2006). "Eliminating Design Alternatives Based on Imprecise Information." Society of Automotive Engineering World Congress, Paper No. 2006-01-0272.

Robbins, H. (1955). "An Empirical Bayes Approach to Statistics." Third Berkeley Symposium on Mathematical Statistics and Probability, 157-63.

Robert, C. P. and G. Casella (1999). *Monte Carlo Statistical Methods*. New York, NY: Springer-Verlag.

Ross, S., D. Evans and M. Webber (2002). "How LCA Studies Deal with Uncertainty." *International Journal of Life Cycle Assessment*, 7(1), 47-52.

Rowe, W. D. (1994). "Understanding Uncertainty." *Risk Analysis*, 14, 743-50.

Rubin, M. R. (1983). *Information Economics and Policy in the U.S.* Littleton, CO: Libraries Unltd.

Sacks, J., W. J. Welch, T. J. Mitchell and H. P. Wynn (1989). "Design and Analysis of Computer Experiments." *Statistical Science*, 4, 409-35.

Sarin, R. K. (1978). "Elicitation of Subjective Probabilities in the Context of Decision-Making." *Decision Sciences*, 9, 37-48.

Savage, L. J. (1972). *The Foundation of Statistics*. New York: Dover Publications.

Schervish, M. J., T. Seidenfeld, J. B. Kadane and I. Levi (2003). *Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons*. Lugano, Switzerland.

Schlesinger, S., R. E. Crosbie, G. S. Innis, C. S. Lalwani, J. Loch, R. J. Sylvester, R. O. Wright, N. Kheir and D. Bartos (1979). "Terminology for Model Credibility." *Simulation*, 32(3), 103-4.

Scott, M. J. (2004). "Utility Methods in Engineering Design." *Engineering Design Reliability Handbook*. New York: CRC Press.

Seepersad, C. C. (2001). *The Utility-Based Compromise Decision Support Problem with Applications in Product Platform Design*. Masters Thesis, Georgia Institute of Technology.

Seidenfeld, T., M. J. Schervish and J. B. Kadane (1995). "A Representation of Partially Ordered Preferences." *The Annals of Statistics*, 23(9), 2168-217.

Sentz, K. and S. Ferson (2002). "Combination of Evidence in the Dempster-Shafer Theory." SAND2002-0835. Sandia National Laboratories, Albuquerque, NM. <http://www.sandia.gov/epistemic/Reports/SAND2002-0835.pdf >.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press.

Shafer, G. (1978). "Non-Additive Probabilities in the Work of Bernoulli and Lambert." *Archive for History of Exact Sciences*, 19, 309-70.

Shafer, G. (1986). "The Combination of Evidence." *International Journal of Intelligent Systems*, 1, 155-79.

Shafer, G. (1992). "Rejoinders to Comments on "Perspectives on the Theory and Practice of Belief Functions"." *International Journal of Approximate Reasoning*, 6(3), 445-80.

Shannon, C. E. (1948). "A Mathematical Theory of Communication." *Bell System Technical Journal*, 27, 379-423, 623-56.

Simon, H. (1947). *Administrative Behavior*. New York: The Macmillan Company.

Simon, H. (1982). *The Sciences of the Artificial (Second Edition)*. Cambridge, MA: The MIT Press.

Simpson, T. W., J. D. Peplinski, P. N. Koch and J. K. Allen (2001). "Metamodels for Computer-Based Engineering Design." *Engineering with Computers*, 17(2), 129-50.

Smets, P. (1990). "The Transferable Belief Model: A Model for Quantified Beliefs." IEE Colloquium on `Reasoning Under Uncertainty' (Digest No.086), 22 May 1990, London, UK, IEE, 4-1.

Smith, C. A. B. (1961). "Consistency in Statistical Inference and Decision." *Journal of Royal Statistical Society*, B23, 1-37.

Sobek, D. K., II, A. C. Ward and J. Liker (1999). "Toyota's Principles of Set-Based Concurrent Engineering." *Sloan Management Review*, 40(2), 67-83.

Sobek, D. K. I. (2004). "Explaining the Second Toyota Paradox by Modeling Real Options in Product Development." *IEEE Transaction on Engineering Management*

Sobek, D. K. I. and A. C. Ward (1996). "Principles from Toyota's Set-Based Concurrent Engineering Process." ASME 1996 International Design Engineering Technical

Conferences and Computers in Engineering Conference, Irvine, California, August 18-22.

Sobol, I. M. (2001). "Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates." *Mathematics and Computers in Simulation*, 55, 271-80.

Stamatelatos, M. (2000). "Probabilistic Risk Assessment: What Is It and Why Is It Worth Performing It?" NASA Office of Safety and Mission Assurance.

Stevens, S. S. (1935). "The Operational Definition of Psychological Concepts." *Psychological Review*, 42, 517-27.

Stevens, S. S. (1936). "Psychology: The Propaedeutic Science." *Philosophy of Science*, 3, 90-103.

Stevens, S. S. (1939). "Psychology and the Science of Science." *Psychological Bulletin*, 36, 221-63.

Strassmann, P. A. (1999). *Information Productivity*. Information Economics Press.

Strassmann, P. A. (2004). "The Foundations of Information Economics - Part I." *Information Economics Journal*, March 2004

Sugeno, M. (1977). "Fuzzy Measures and Fuzzy Integrals: A Survey." In M. M. Gupta, G. N. Saridis and B. R. Gaines (Eds.), *Fuzzy Automata and Decision Processes*.

Szykman, S., R. Sriram and W. C. Regli (2001). "The Role of Knowledge in Next-Generation Product Development Systems." *ASME Journal of Computing and Information Science in Engineering*, 1(1), 3-11.

Taguchi, G. (1987). *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*. (L. W. Tung. Trans.). Dearborn, Michigan: UNIPUB/Kraus International Publications.

Tesfasion, L. (1976). "Stochastic Dominance and the Maximization of Expected Utility." *Review of Economic Studies*, 43, 301-15.

Thurston, D. L. (1990a). "Multiattribute Utility Analysis in Design Management." *IEEE Transactions on Engineering Management*, 37(4), 296-301.

Thurston, D. L. (1990b). "Subjective Design Evaluation with Multiple Attributes." 1990 ASME Design Technical Conferences - Presented at the 2nd International Conference on

Design Theory and Methodology, Sep 16-19 1990, Chicago, IL, USA, Publ by ASME, New York, NY, USA, 355-61.

Thurston, D. L. (2001). "Real and Misconceived Limitations to Decision Based Design with Utility Analysis." *Journal of Mechanical Design*, 123, 176-82.

Thurston, D. L., J. V. Carnahan and T. Liu (1991). "Optimization of Design Utility." 3rd International Conference on Design Theory and Methodology presented at the 1991 ASME Design Technical Conferences, Sep 22-25 1991, Miami, FL, USA, Publ by ASME, New York, NY, USA, 173-80.

Tintner, G. (1941). "The Theory of Choice under Subjective Risk and Uncertainty." *Econometrica*, 9, 298-304.

Tong, C. and D. Sriram (1992). *Artificial Intelligence in Engineering Design: Models of Innovative Design, Reasoning About Physical Systems, and Reasoning About Geometry (Artificial Intelligence in Engineering Design).* Boston: Academic Press.

Trejo, R. and V. Kreinovich (2001). "Error Estimations for Indirect Measurements: Randomized Vs. Deterministic Algorithms for 'Black-Box' Programs." In S. Rajasekaran, P. Pardalos, J. Reif and J. Rolim (Eds.), *Handbook on Randomized Computing* (pp. 673-729). Dordrecht: Kluwer.

Tucker, W. T. and S. Ferson (2003). "Probability Bounds Analysis in Environmental Risk Assessment." Applied Biomathematics, Setauket, New York. <http://www.ramas.com/pbawhite.pdf >.

Tversky, A. and D. Kahneman (1974). "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185, 1124-31.

Urban, S. D., S. W. Dietrich, A. Saxena and A. Sundermier (2001). "Interconnection of Distributed Components: Current Middleware Solutions." *ASME Journal of Computing and Information Science in Engineering*, 1(1)

Utkin, L. V. (2003). "A Second-Order Uncertainty Model of Independent Random Variables: An Example of the Stress-Strength Reliability." Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications, Lugano, Switzerland.

Utkin, L. V. and T. Augustin (2003). "Decision Making with Imprecise Second-Order Probabilities." Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications, Lugano, Switzerland.

von Fraassen, B. (1989). *Laws and Symmetry*. Oxford: Clarendon Press.

von Mises, R. (1939). *Probability, Statistics and Truth*. (Second ed.). New York: Macmillan.

von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

von Neumann, J. and O. Morgenstern (1980). *Theory of Games and Economic Behavior, Third Edition*. Princeton, NJ: Princeton University Press.

Vose, D. (2000). *Risk Analysis, a Quantitative Guide*. (2nd Edition ed.). New York: John Wiley and Sons.

Walley, P. (1981). "Coherent Lower (and Upper) Probabilities." Statistics Research Report 22. University of Warwick, Coventry.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. New York: Chapman and Hall.

Walley, P. (1996). "Measures of Uncertainty in Expert Systems." *Artificial Intelligence*, 83(1), 1-58.

Walley, P. (1997). "Statistical Inferences Based on a Second-Order Possibility Distribution." *International Journal of General Systems*, 26(4), 337-83.

Walley, P. (2000). "Towards a Unified Theory of Imprecise Probability." *International Journal of Approximate Reasoning*, 24(2-3), 125-48.

Walley, P. and G. De Cooman (2001). "A Behavioral Model for Linguistic Uncertainty." *Information Sciences*, 134(1-4), 1.

Ward, A. C., J. K. Liker, J. J. Cristiano and D. K. Sobek II (1995). "The Second Toyota Paradox: How Delaying Decisions Can Make Better Cars Faster." *Sloan Management Review*, 43-61.

Watson, S. R., J. J. Weiss and M. L. Donnell (1979). "Fuzzy Decision Analysis." *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 1-9.

Weber, M. (1987). "Decision Making with Incomplete Information." *European Journal of Operational Research*, 28(1), 44.

Weichselberger, K. (2000). "The Theory of Interval-Probability as a Unifying Concept for Uncertainty." *International Journal of Approximate Reasoning*, 24(2-3), 149.

Williamson, R. C. (1989). *Probabilistic Arithmetic*. Ph.D. Thesis, University of Queensland.

Williamson, R. C. and T. Downs (1990). "Probabilistic Arithmetic I: Numerical Methods for Calculating Convolutions and Dependency Bounds"." *International Journal of Approximate Reasoning*, 4, 89-158.

Winkler, R. L. (1996). "Uncertainty in Probabilistic Risk Assessment." *Reliability Engineering & System Safety*, 54(2-3), 127-32.

Yager, R. R. (1986). "Arithmetic and Other Operations on Dempster-Shafer Structures." *International Journal of Man-Machine Studies*, 25, 357-66.

Yager, R. R., M. Fedrizzi and J. Kacprzyk (Eds.) (1994a). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley and Sons, Inc.

Yager, R. R., J. Kacprzyk and M. Fedrizzi (1994b). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley and Sons.

Zadeh, L. A. (1965). "Fuzzy Sets." *Information and Control*, 8, 338-53.

Zadeh, L. A. (1978). "Fuzzy Sets as a Basis for a Theory of Possibility." *Fuzzy Sets and Systems*, 1, 3-28.

Zadeh, L. A. (1984a). "Fuzzy Probabilities." *Information Processing and Management*, 20, 363-72.

Zadeh, L. A. (1984b). "Review of Books: A Mathematical Theory of Evidence." *The AI Magazine*, 5(3), 81-3.

Zaffalon, M., K. Wesnes and O. Petrini (2003). "Reliable Diagnoses of Dementia by the Naive Credal Classifier Inferred from Incomplete Cognitive Data." *Artificial Intelligence in Medicine*, 29(1-2), 61-79.

Zimmerman, H. J. (1996). *Fuzzy Set Theory*. Norwell, MA: Kluwer Academic Publishers.